

Chapter

7

More about Statistical Graphs

Learning Objectives

After completing this chapter, you will be able to

- construct and interpret frequency polygons and frequency curves.
- construct and interpret cumulative frequency polygons and cumulative frequency curves.
- read the data associated with percentiles and quartiles from the graphs.



1



2

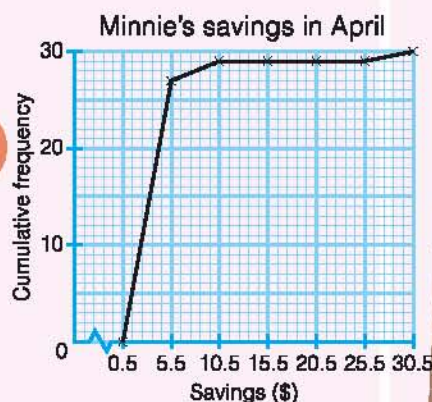


3



4

Minnie's savings in April is shown in the following figure. She thinks that her total savings in that month will be less than \$210. Do you agree with Minnie? Explain briefly.





Preview

[Basic knowledge required for this chapter.]

Basic Knowledge

1. Concepts about classes

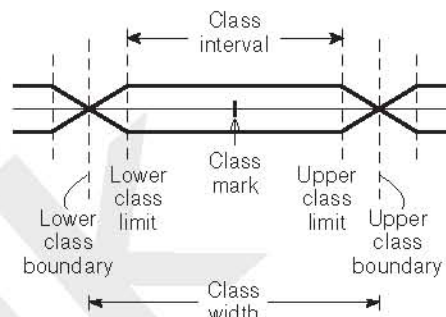
e.g. For the class interval 21 cm - 30 cm (correct to the nearest cm),

lower class limit = 21 cm, upper class limit = 30 cm;

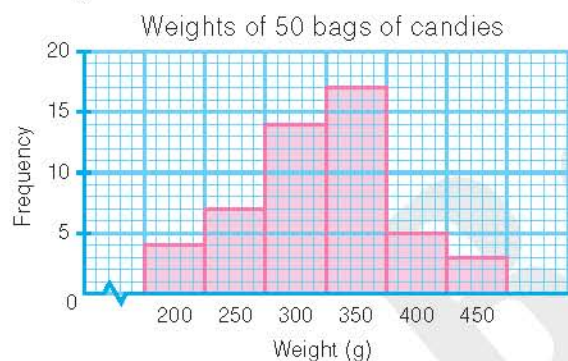
lower class boundary = 20.5 cm, upper class boundary = 30.5 cm;

class width = $(30.5 - 20.5)$ cm = 10 cm;

class mark = $\frac{21 + 30}{2}$ cm = 25.5 cm.



2. Histogram



7.1 Frequency Polygons and Frequency Curves

Apart from using a histogram, we can also use the statistical graphs introduced in this section to present and analyze the distribution of continuous data.

A Construction of frequency polygons and frequency curves

To investigate the customer flows, Hamburg Restaurant collected the time spent by 100 customers in the restaurant, and presented the data with a frequency distribution table as follows.

Time spent in the restaurant (class width = 10 min)

Time (min)	21 - 30	31 - 40	41 - 50	51 - 60	61 - 70	71 - 80
Class boundaries (min)	20.5 - 30.5	30.5 - 40.5	40.5 - 50.5	50.5 - 60.5	60.5 - 70.5	70.5 - 80.5
Class mark (min)	25.5	35.5	45.5	55.5	65.5	75.5
Frequency	18	24	33	13	8	4

Table 7.1

Based on the distribution of data above, the following histogram can be obtained.

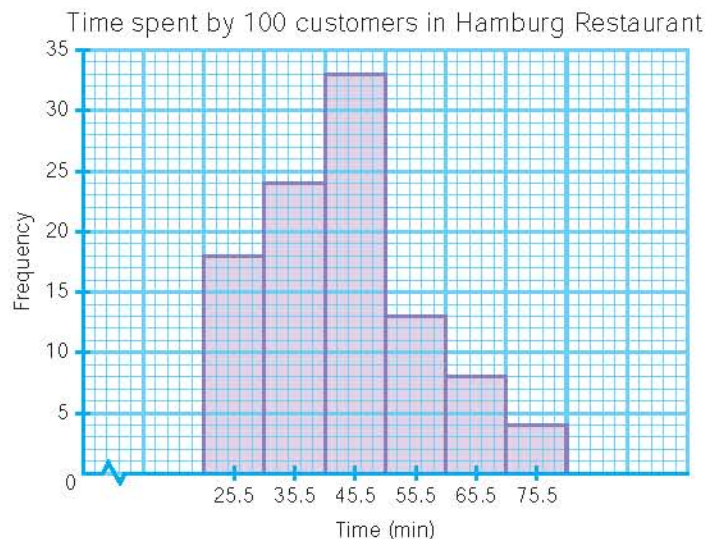


Figure 7.1

At the two ends of the histogram, add a class interval each with zero frequency. Then join the mid-points at the top of the bars (including mid-points of the two additional class intervals) and the following graph is obtained.

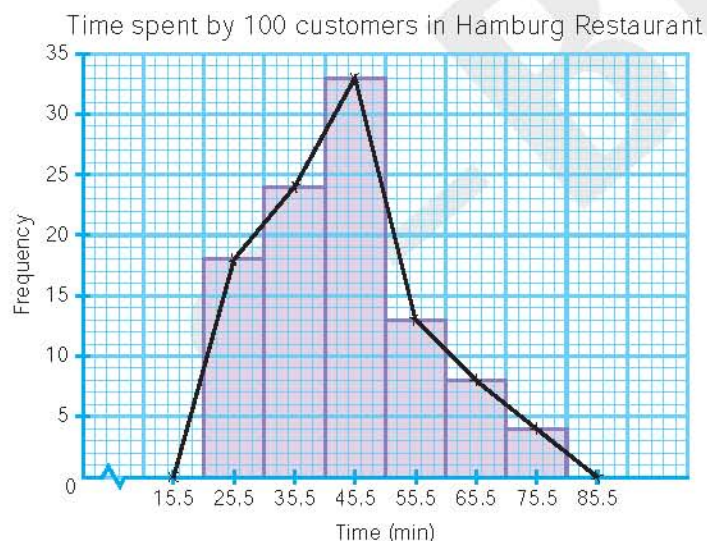


Figure 7.2

For a set of data, there are many different frequency polygons depending on how the data are grouped.



The above graph is known as a **frequency polygon**, which can display the distribution of the data.

Obviously, the mid-points at the top of the bars can be directly marked according to the class marks and their corresponding frequencies in the frequency distribution table, and hence a frequency polygon can be obtained.

The histogram is not required to be drawn before constructing a frequency polygon, but all the points (where '×' is used in this book) must be marked clearly.

frequency polygon 頻數多邊形

A **frequency curve** can be drawn by smoothing out the frequency polygon.

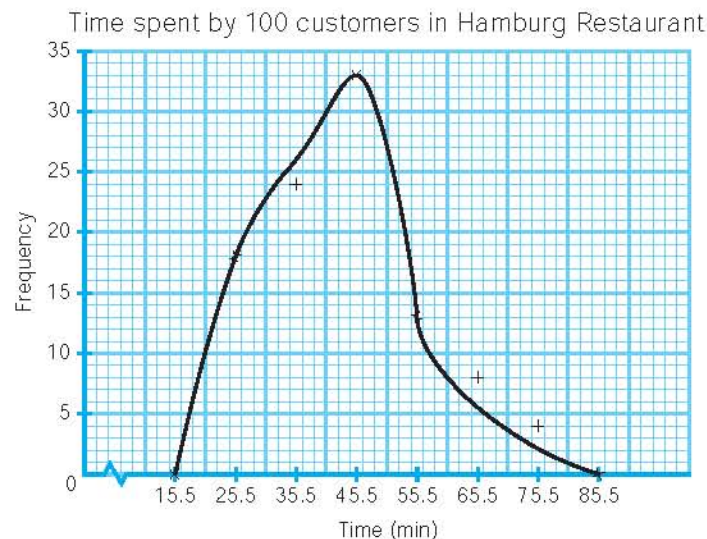


Figure 7.3

Frequency curve only reflects the outline of the distribution, and thus, it does not necessarily pass through all the marked points.



Extension 7.1

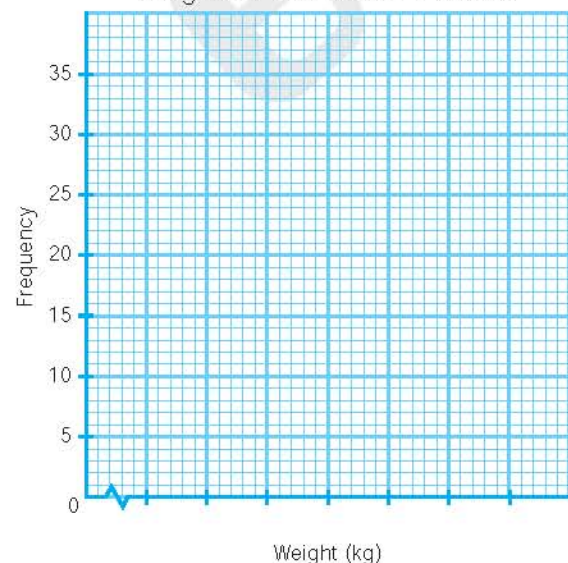
Kindness Hospital records the weights of 100 newborn babies in the following table.

Weight (kg)	2.1 - 2.5	2.6 - 3.0	3.1 - 3.5	3.6 - 4.0	4.1 - 4.5
Class mark (kg)	2.3	2.8	3.3	3.8	4.3
Frequency	6	19	34	31	10



Draw a frequency polygon to represent the weight distribution of the babies above.

Weights of 100 newborn babies



frequency curve 頻數曲綫

B Understanding frequency polygons and frequency curves

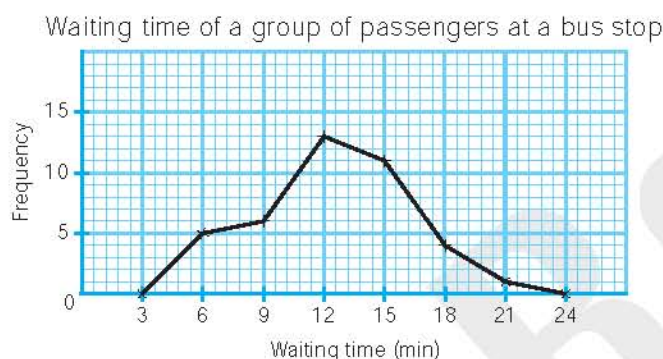
I. Reading frequencies

From a marked point of a frequency polygon, we can obtain the frequency corresponding to a class interval. From the line segments joining two adjacent points, we can see the general distribution of the data.



Example 7.1 Reading information from a frequency polygon

The following frequency polygon shows the waiting time of a group of passengers at a bus stop.



- What is the class width?
- Based on the frequency polygon, complete the following table.

Waiting time (min)	Class mark (min)	Frequency
5 - 7	6	
8 - 10	9	
Total		

- Which class interval do most passengers belong to? How many passengers are there?
- How many passengers are there in the group?
- What percentage of the passengers have waited for less than 13.5 min?

Solution

- (a) Class width = $(6 - 3)$ min
 $= \underline{\underline{3 \text{ min}}}$

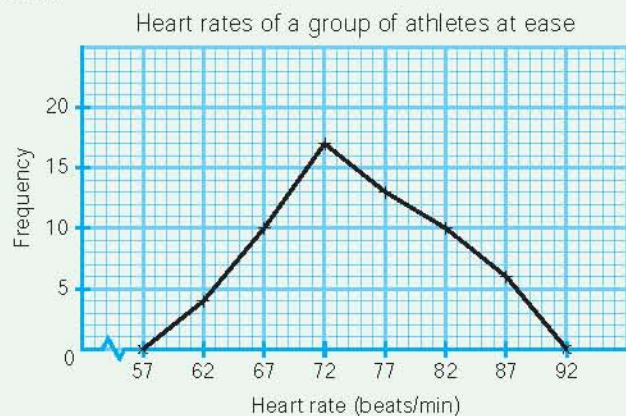
(b)

Waiting time (min)	Class mark (min)	Frequency
5 - 7	6	5
8 - 10	9	6
11 - 13	12	13
14 - 16	15	11
17 - 19	18	4
20 - 22	21	1
Total		40

- (c) Most passengers belong to the class interval 11 min - 13 min, and there are 13 passengers.
- (d) There are 40 passengers in the group.
- (e) The number of the passengers who have waited for less than 13.5 min
 $= 5 + 6 + 13$
 $= 24$
 \therefore The required percentage $= \frac{24}{40} \times 100\%$
 $= \underline{\underline{60\%}}$

Classwork 7.1

The following frequency polygon shows the heart rates of a group of athletes at ease.



- | Heart rate (beats/min) | Class mark (beats/min) | Frequency |
|------------------------|------------------------|-----------|
| 60 - 64 | 62 | |
| 65 - 69 | 67 | |
| | | |
| | | |
| | | |
| | | |
| | | |
| Total | | |

- 
- Example

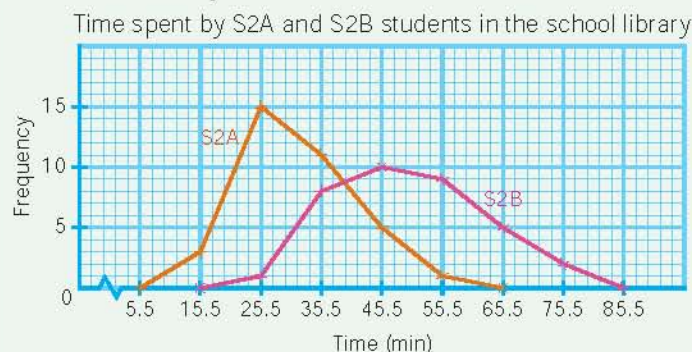
Solution

-
- The graph displays the frequency of scores for two groups, S2A and S2B. The x-axis represents the score, ranging from 0 to 75.5 with major grid lines every 10 units and minor grid lines every 5 units. The y-axis represents the frequency, ranging from 0 to 15 with major grid lines every 5 units and minor grid lines every 1 unit. S2A is represented by an orange line, and S2B is represented by a purple line. S2A starts at a frequency of 0 at a score of 5.5, rises to a peak frequency of 15 at a score of 45.5, and then falls to 0 at a score of 65.5. S2B starts at a frequency of 0 at a score of 15.5, rises to a peak frequency of 14 at a score of 55.5, and then falls to 0 at a score of 75.5.
- | Score | S2A Frequency | S2B Frequency |
|-------|---------------|---------------|
| 5.5 | 0 | 0 |
| 15.5 | 3 | 0 |
| 25.5 | 6 | 3 |
| 35.5 | 12 | 9 |
| 45.5 | 15 | 11 |
| 55.5 | 4 | 14 |
| 65.5 | 0 | 3 |
| 75.5 | 0 | 0 |

- (b) Considering the students whose scores are 50.5 or above,
 number of students in S2A = 4
 number of students in S2B = $14 + 3$
 $= 17$
- (c) From (a) and (b), we can see that more students in S2A have got lower scores, and less students there have got higher scores. Thus the overall performance of S2A students is worse than that of S2B students.

Classwork 7.2

The following frequency polygons show the time spent by S2A and S2B students in the school library.

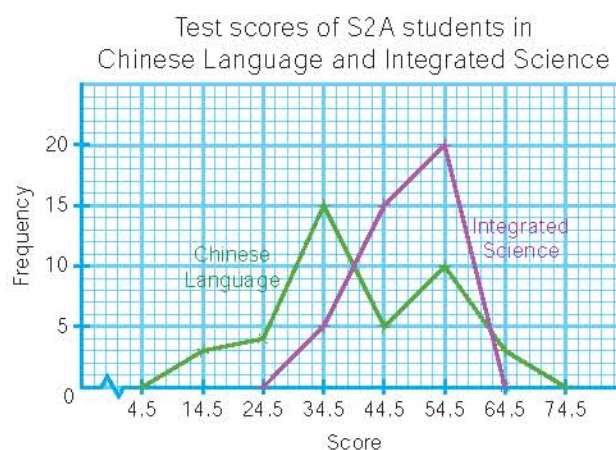


- (a) How many students are there in each class who have spent less than 30.5 min in the library?
 (b) How many students are there in each class who have spent 50.5 min or more in the library?
 (c) In general, which class of students have spent more time in the library?

Skills Upgrading Corner 7.1

The frequency polygons show the test scores of S2A students in Chinese Language and Integrated Science.

- (a) In which subject did students get scores below 19.5?
 (b) In which subject did students get scores closer to each other?
 (c) If the pass mark is 40, find the percentages of students passing the tests in Chinese Language and Integrated Science each.
 (d) In which subject did students get a better overall performance?



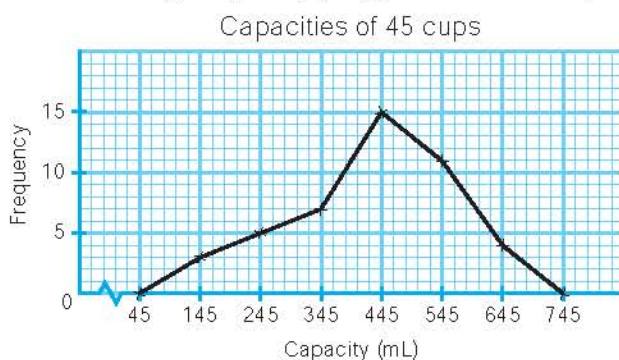


Exercise 7A

[Graph paper is provided in the Appendix.]

Level 1

1. The following frequency polygon shows the capacities of 45 cups.



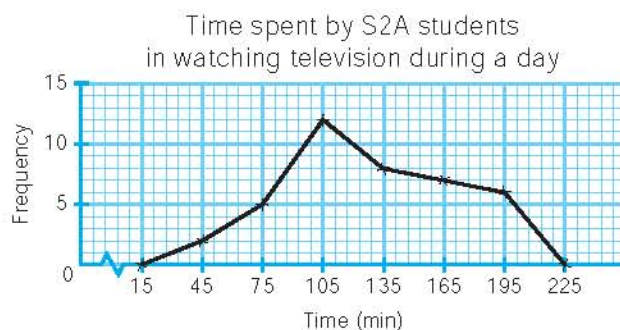
- (a) Based on the frequency polygon, complete the following table.

Capacity (mL)	100 - 190	200 - 290				
Class mark (mL)	145	245				
Frequency						

- (b) How many cups are there in the class interval 200 mL - 290 mL?
 (c) Which class interval has most of the cups? How many cups are there?

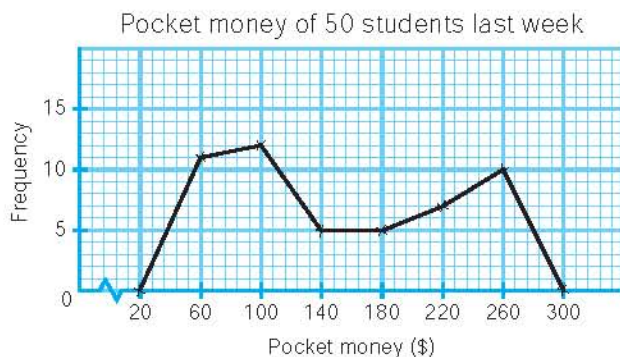
2. The frequency polygon shows the time spent by S2A students in watching television during a day.

- (a) What is the class width?
 (b) How many students are there in this class?
 (c) How many students spent about 1 to 2 hours in watching television?
 (d) Do you agree that students in S2A like watching television? Why?



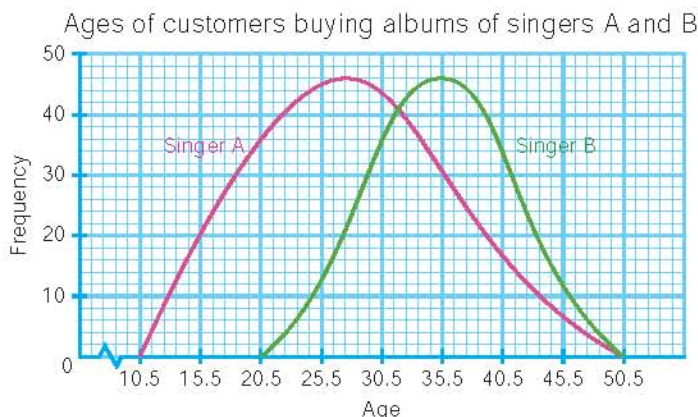
3. The frequency polygon shows the amount of pocket money of 50 students last week.

- (a) How many class intervals are there? What is the class width?
 (b) How many students had pocket money less than \$200 last week?
 (c) How many students had pocket money equal to \$280 or above last week?
 (d) What percentage of the students had pocket money less than \$120?

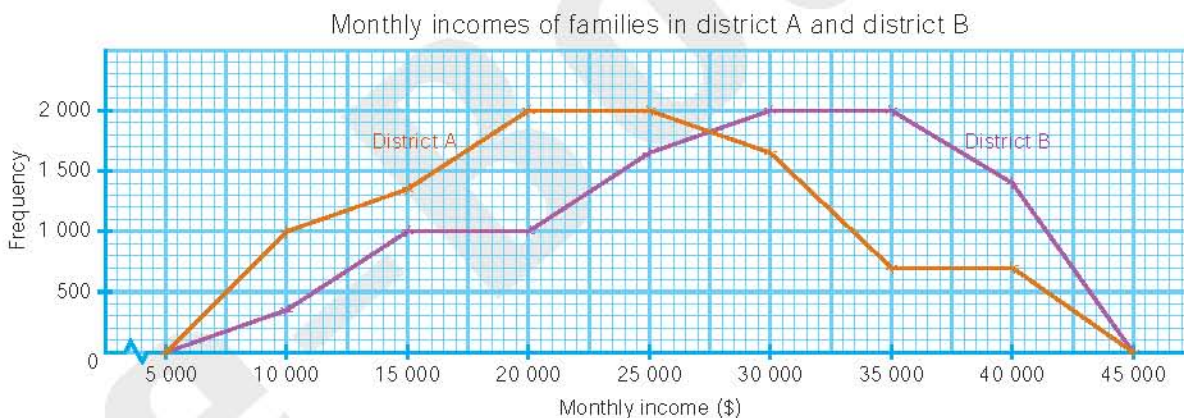


Level 2

4. The following frequency curves show the ages of customers buying albums of singers A and B in an audio store yesterday.



- Which album attracts customers younger than 18?
 - Which singer has younger fans?
5. The following frequency polygons show the monthly incomes of families in district A and district B.



- Which district has families with lower monthly income?
 - If families with their monthly incomes less than \$12 500 are classified as low income families, which district has more low income families?
 - If families with their monthly incomes equal to \$32 500 or above are classified as high income families, which district has more high income families?
 - According to the above figure, Sally believes that there are 3 000 families in total from district A and district B which earn \$20 000 a month. Do you agree with Sally? Why?
6. The following are the scores of a basketball team in 40 matches.

73	64	78	66	52	88	48	72	58	51	65	85	66	69	79	63	81	84	53	68
86	76	69	54	67	76	54	59	74	62	44	76	53	62	86	60	51	67	53	67

- Construct a frequency distribution table using 40 - 49 as the first class interval.
- Draw a frequency polygon.

7. The following table shows the distribution of the test scores of 50 students in Mathematics and Chinese Language (the full marks are 100).

Class mark		14.5	24.5	34.5	44.5	54.5	64.5	74.5	84.5
Frequency	Mathematics	11	10	14	7	0	3	3	2
	Chinese Language	2	4	0	0	2	13	16	13

- (a) Draw frequency polygons on the same figure to present the scores of the two subjects.
- (b) What are the characteristics of the performances of the students in the two subjects?
8. The following frequency distribution table shows the lifetime of light bulbs (in hours) produced by production lines A and B.

Lifetime (h)	100 - 199	200 - 299	300 - 399	400 - 499	500 - 599
Production line A	9	17	35	20	19
Production line B	13	36	31	12	8

- (a) Draw frequency polygons on the same figure to show the lifetime of light bulbs produced by production lines A and B.
- (b) Which production line produces light bulbs with longer lifetime?
- (c) If a light bulb produced by production line A is chosen, which class interval does its lifetime most possibly fall into?

7.2 Cumulative Frequency Polygons and Cumulative Frequency Curves

Class Activity 7.1

Aim: To learn the concept of cumulative quantity

1. Alex saved \$100 in total from Monday to Friday. Figure I shows the cumulative savings of him for these 5 days.

- (a) Up to Tuesday, the total amount saved by Alex was \$ 40.
- (b) The amount saved by Alex on Wednesday was \$ 20.
- (c) Alex saved most money on Tuesday and he saved \$ 30.



2. The following figures show the cumulative savings of Frankie and Kathy for the past 5 days. They wanted to save \$100 on or before Friday.

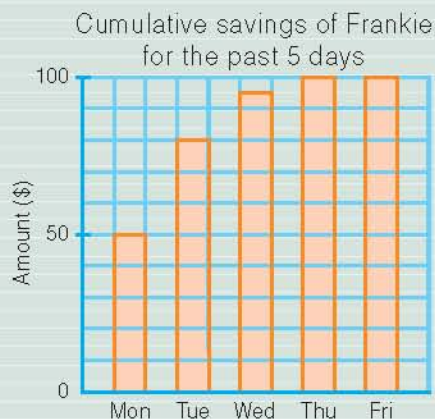


Figure II

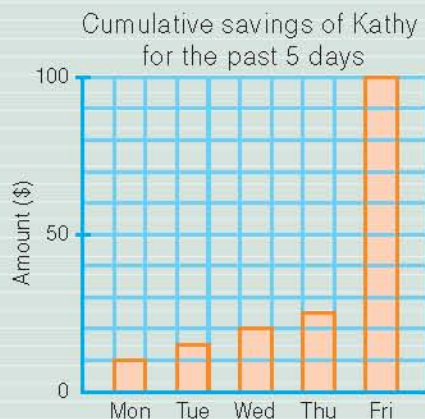


Figure III

- (a) From the figures, describe their saving patterns.

Frankie saved most of the money from Monday to Wednesday.

Kathy saved most of the money on Friday.

- (b) Who saved \$100 earlier?



Frankie



Kathy

Now I see ...

We can use diagrams with cumulative quantities to show the increase in quantities of an object.



In addition to the quantities of each class of grouped data, the cumulative quantities of a set of grouped data are often considered in statistics.

The weights (correct to the nearest kg) of 40 students are listed in Figure 7.4.

We can present the data with a frequency distribution table as follows:

Weight (kg)	Class boundaries (kg)	Frequency
46 - 50	45.5 - 50.5	2
51 - 55	50.5 - 55.5	6
56 - 60	55.5 - 60.5	9
61 - 65	60.5 - 65.5	12
66 - 70	65.5 - 70.5	8
71 - 75	70.5 - 75.5	3

Table 7.2

74	46	71	65	56
54	67	54	61	62
66	52	68	70	59
49	68	64	62	75
51	57	53	57	60
69	60	70	63	65
63	60	54	59	62
67	60	62	62	61

Figure 7.4

From Table 7.2, we know that there are no students weighing less than 45.5 kg, 2 students weighing less than 50.5 kg, 8 students weighing less than 55.5 kg and so on. The number of students (2, 8, ...) obtained above is called **cumulative frequency**. The following **cumulative frequency table** can be constructed to represent the above information.

Number of students weighing less than 55.5 kg = $2 + 6 = 8$.

Weight less than (kg)	Frequency	Cumulative frequency
45.5	0	0
50.5	2	2
55.5	6	8
60.5	9	17
65.5	12	29
70.5	8	37
75.5	3	40

Being the ordered pairs

Table 7.3

If we mark the ordered pairs (45.5, 0), (50.5, 2), (55.5, 8), ..., (75.5, 40) on a figure and join all the points, the following **cumulative frequency polygon** can be obtained.

Weights of 40 students

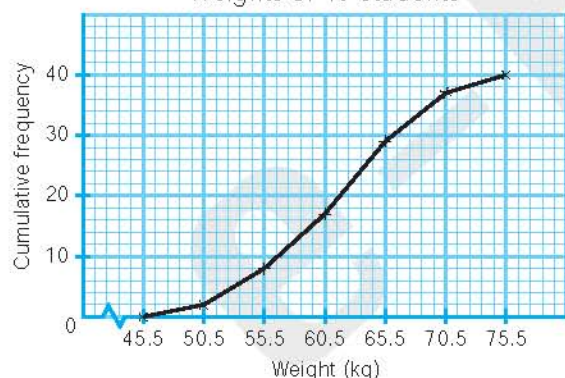


Figure 7.5

In a cumulative frequency polygon, each line segment would never go downwards.



Example 7.3 Constructing cumulative frequency polygon

The following table shows the distribution of ages of 50 employees in Hugo Company.

Age	21 - 25	26 - 30	31 - 35	36 - 40	41 - 45	46 - 50
Frequency	6	16	10	9	5	4

- Construct a cumulative frequency table for the above data.
- Draw a cumulative frequency polygon.

cumulative frequency 累積頻數

cumulative frequency table 累積頻數表

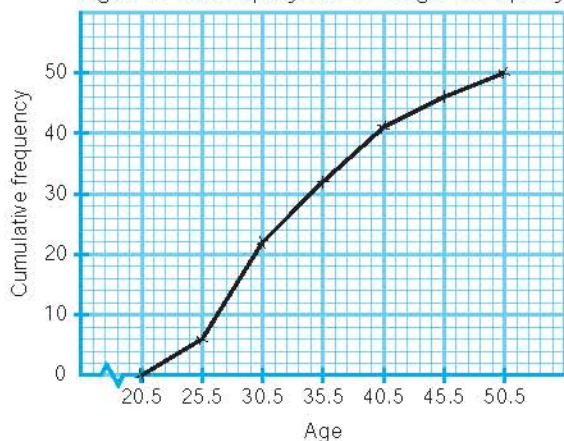
cumulative frequency polygon 累積頻數多邊形

Solution

(a)

Age less than	20.5	25.5	30.5	35.5	40.5	45.5	50.5
Cumulative frequency	0	6	22	32	41	46	50

(b) Ages of 50 employees in Hugo Company



Classwork 7.3

The following table shows the time spent by S2D students on physical exercise per week.

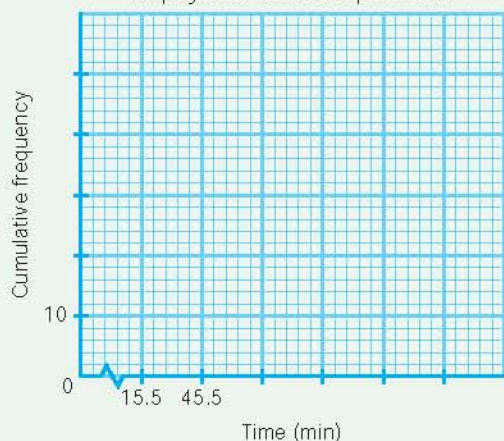
Time (min)	16 - 45	46 - 75	76 - 105	106 - 135	136 - 165
Frequency	6	12	14	8	2

(a) Complete the following table.

Time less than (min)	15.5	45.5	75.5	105.5	135.5	165.5
Cumulative frequency						

(b) Draw a cumulative frequency polygon.

Time spent by S2D students
on physical exercise per week



The cumulative frequency polygon in Figure 7.5 shows the weight distribution of 40 students. We can smooth out a cumulative frequency polygon to form a **cumulative frequency curve**. (See Figure 7.6)

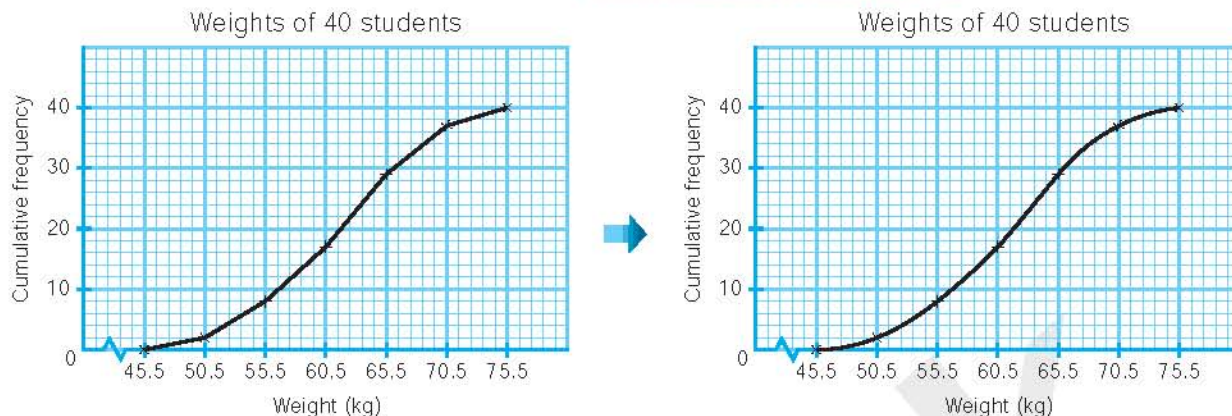


Figure 7.6



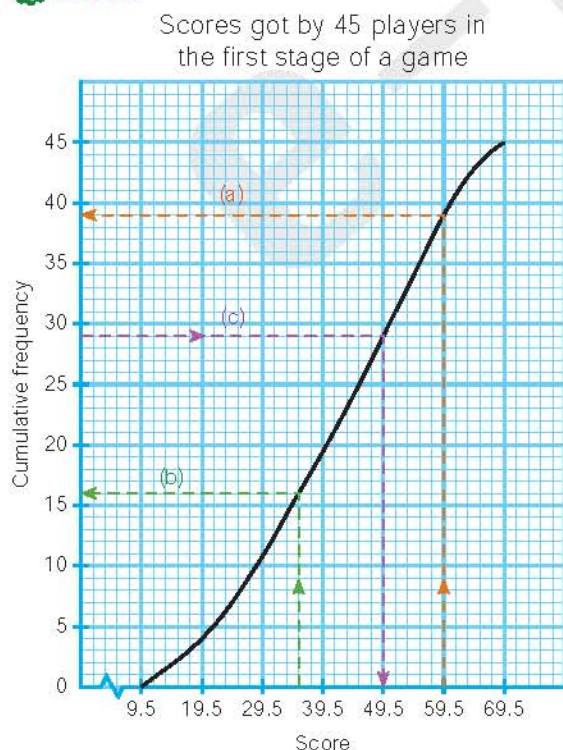
[Refer to page 7.42]

Example 7.4 Getting information from cumulative frequency curve

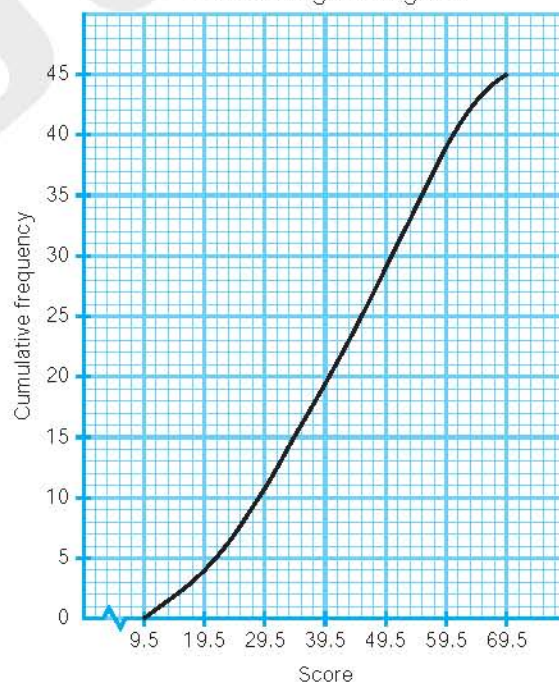
The cumulative frequency curve shows the scores got by 45 players in the first stage of a game.

- Find the number of players whose scores are below 59.5.
- Find the number of players whose scores are 35.5 or above.
- If the 16 players with the highest scores can get into the second stage of the game, what is the minimum score to get into the second stage?

Solution



Scores got by 45 players in the first stage of a game



- (a) From the graph,
number of players whose scores are below 59.5 = 39
- (b) From the graph,
number of players whose scores are below 35.5 = 16
 \therefore Number of players whose scores are 35.5 or above
 $= 45 - 16$
 $= \underline{29}$

- (c) [Analysis: If only 16 players can get into the second stage, there are
 $(45 - 16) = 29$ players who fail to get into the second stage.]

From the graph,

when the cumulative frequency = 29,

the corresponding score = 49.5.

\therefore The minimum score to get into the second stage = 49.5

Draw a vertical line at the score of 59.5 from the horizontal axis to the curve. At the intersection point on the curve, draw a horizontal line intersecting the vertical axis. The reading of 39 from the vertical axis means 39 players got less than the score of 59.5.

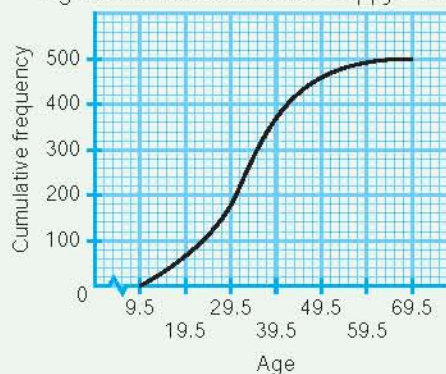


Classwork 7.4

The cumulative frequency curve shows the ages of 500 customers of Happy Fast Food.

- (a) Find the number of customers whose ages are below 29.5.
 (b) Find the number of customers whose ages are 37.5 or above.
 (c) If the fast food shop let 40 eldest customers line up first, find the age of the youngest one among them.

Ages of 500 customers of Happy Fast Food

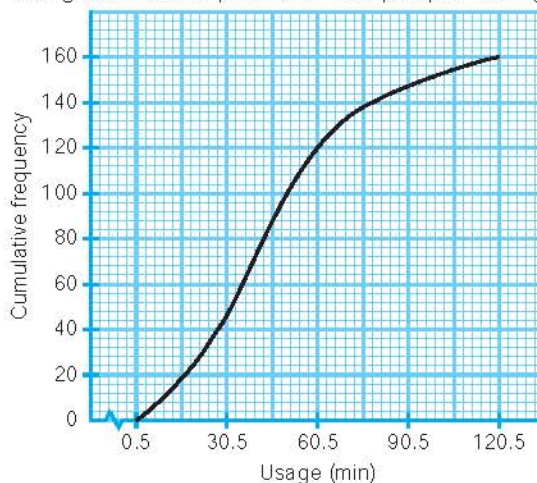


Example 7.5 Finding the required information by using cumulative frequency curve

The cumulative frequency curve shows the usage of mobile phone of 160 people during a day.

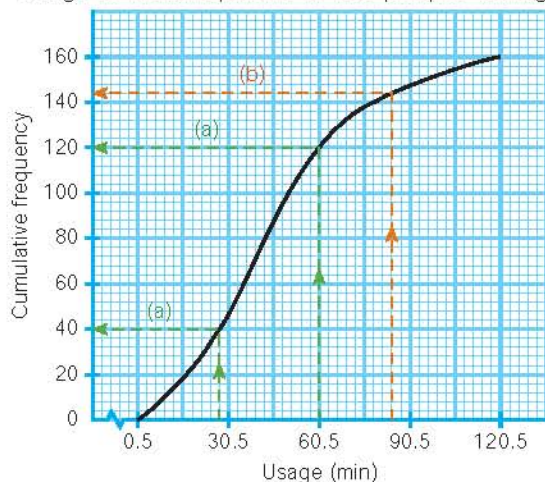
- (a) Find the number of people whose usage is between 27.5 min and 60.5 min.
 (b) What percentage of people whose usage is less than 84.5 min?

Usage of mobile phone of 160 people during a day



Solution

Usage of mobile phone of 160 people during a day



- (a) [Analysis: From the graph, there are 40 people whose usage is less than 27.5 min, and there are 120 people whose usage is less than 60.5 min.]

$$\begin{aligned} \text{Number of people whose usage is between 27.5 min and 60.5 min} &= 120 - 40 \\ &= \underline{80} \end{aligned}$$

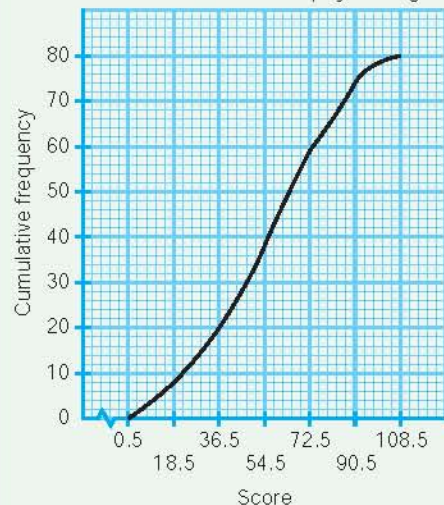
- (b) [Analysis: From the graph, there are 144 people whose usage is less than 84.5 min.]

$$\begin{aligned} \text{The required percentage} &= \frac{144}{160} \times 100\% \\ &= \underline{90\%} \end{aligned}$$

Classwork 7.5

The cumulative frequency curve shows the scores of 80 students in a psychological test about self confidence.

Scores of 80 students in a psychological test



- (a) Find the number of students whose scores are 45.5 or above.
 (b) Any students who have scored between 54.5 and 81.5 are regarded as self-confident. What percentage of students are self-confident?

Skills Upgrading Corner 7.2

The figure shows two cumulative frequency polygons for the scores of a movie given by two groups of audience.

- How many audience are there in each group?
- How many audience in each group gave the scores between 7 and 9?
- Find the percentage of audience in each group whose scores are 6 or above.



Exercise 7B

[Graph paper is provided in the Appendix.]

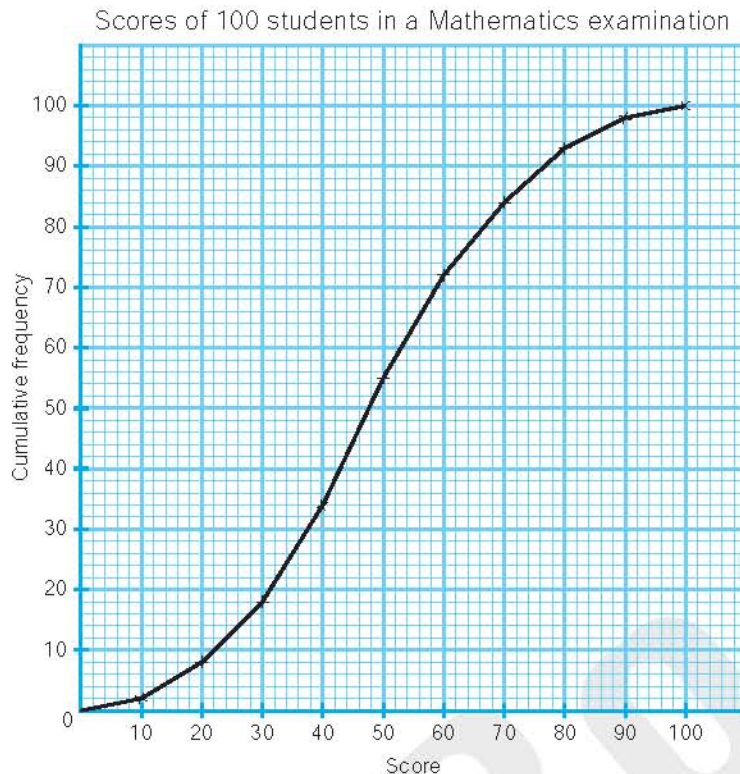
Level 1

- The following cumulative frequency polygon shows the parking hours of cars in a car park during a day.

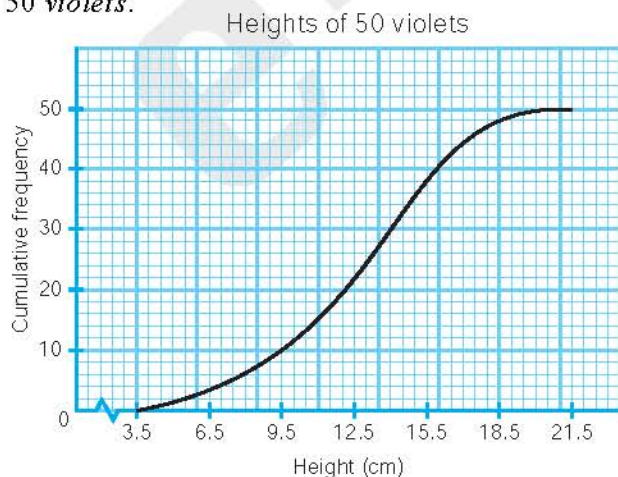


- How many cars were there in the car park?
- How many cars were there with their parking hours less than 5.5 hours?
- How many cars were there with their parking hours of 14 hours or above?

2. The following cumulative frequency polygon shows the scores of 100 students in a Mathematics examination.



- How many students are scoring below 26?
 - How many students are scoring 54 or above?
 - If 10 students with the highest scores can get grade A, what is the minimum score to get a grade A?
3. The following cumulative frequency curve shows the heights of 50 violets.



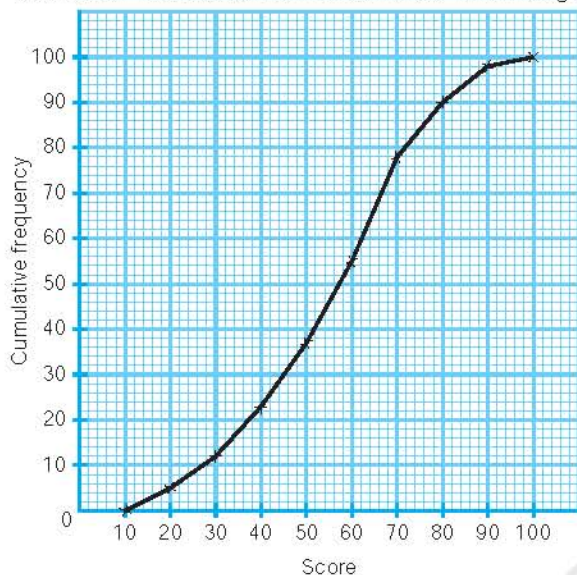
- Find the number of violets whose heights are between 9.5 cm and 12.5 cm.
- If violets taller than or equal to 14 cm should be transferred to larger pots, find the number of violets which is necessary to be transferred.
- If the tallest 40 violets are sold, find the height of the shortest one among them.

violet 紫蘿蘭

Level 2

4. The following cumulative frequency polygon shows the scores of 100 students in a common knowledge quiz.

Scores of 100 students in a common knowledge quiz



- (a) If the top 40% of students will receive a gift each, find the score required for receiving a gift.
- (b) If a student with the score of 80 or above will further receive a certificate, what percentage of students will receive a certificate?
5. The following table shows the distribution of the number of hamburgers sold in a fast food shop for the past 100 days.

Number of hamburgers	551 - 600	601 - 650	651 - 700	701 - 750	751 - 800	801 - 850
Frequency	7	11	34	22	19	7

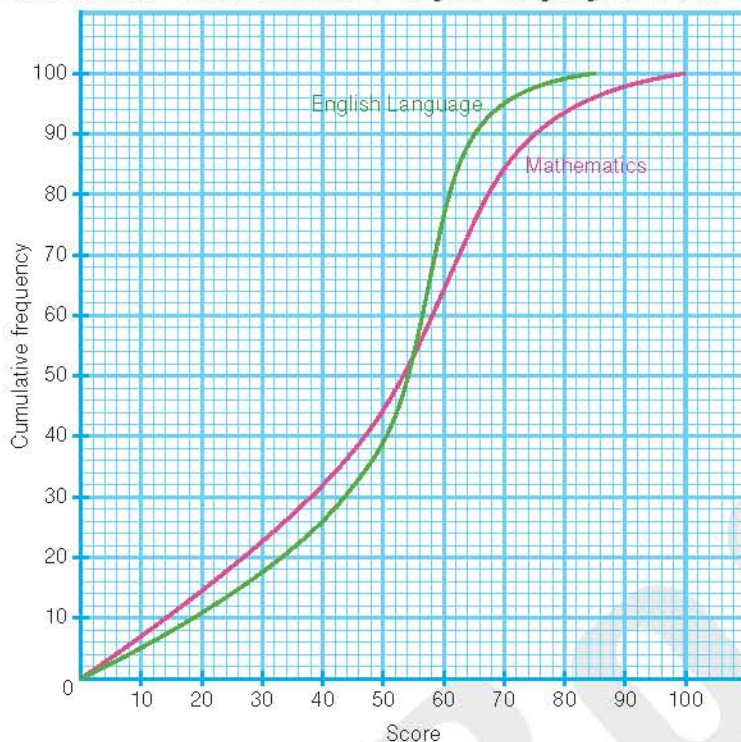
- (a) Construct a cumulative frequency table.
- (b) Draw a cumulative frequency polygon.
[Horizontal scale: 5 squares = 50 hamburgers; vertical scale: 5 squares = 10 days.]
6. The following table shows the monthly salaries of 160 workers.

Salary (\$)	5 500 - 5 990	6 000 - 6 490	6 500 - 6 990	7 000 - 7 490	7 500 - 7 990	8 000 - 8 490	8 500 - 8 990
Frequency	46	28	50	18	10	6	2

- (a) Construct a cumulative frequency table.
- (b) Draw a cumulative frequency polygon.
[Horizontal scale: 10 squares = \$500; vertical scale: 5 squares = 10 workers.]
- (c) From the graph in (b),
- find the number of workers whose monthly salaries are below \$6 795.
 - find the percentage of workers whose monthly salaries are \$7 595 or above.
 - what is the minimum monthly salary for the top 100 income earners?

7. The cumulative frequency curves show the test scores of 100 S2 students in English Language and Mathematics.

Test scores of 100 S2 students in English Language and Mathematics



- (a) What is the highest score in each of the following subjects?
 - (i) English Language
 - (ii) Mathematics
 - (b) How many students scored between 40 and 60 in each of the following subjects?
 - (i) English Language
 - (ii) Mathematics
 - (c) Find the percentage of students whose scores are 70 or above in each of the following subjects.
 - (i) English Language
 - (ii) Mathematics
8. The following shows the lengths (in cm) of 60 babies.

45.8	45.6	46.1	44.9	45.3	45.2	45.7	45.7	44.4	44.8
45.9	44.7	45.1	45.8	46.2	45.9	46.0	46.1	44.5	45.4
45.0	45.3	44.9	45.4	45.4	45.5	45.4	46.3	45.8	45.5
44.7	45.6	46.4	45.7	45.2	45.3	46.6	45.0	46.7	45.5
45.5	46.0	44.6	45.9	45.0	46.0	45.1	45.2	45.6	45.4
45.5	45.6	46.2	45.2	45.8	45.4	45.7	46.5	45.8	45.1



- (a) Construct a frequency distribution table using 44.4 cm - 44.6 cm as the first class interval.
- (b) Construct a cumulative frequency table.
- (c) Draw a cumulative frequency polygon.
[Horizontal scale: 15 squares = 0.3 cm; vertical scale: 10 squares = 10 babies.]
- (d) From the graph in (c),
 - (i) find the number of babies whose lengths are 45.6 cm or above.
 - (ii) find the percentage of babies whose lengths are between 45.1 cm and 46.0 cm.

7.3 Percentiles

A Concept of percentiles

Class Activity 7.2

Aim: To learn the concept of percentiles

The cumulative frequency curve in Figure I shows the average examination scores of S2 students.

(a) The number of S2 students is 200.

(b) Referring to Figure I, complete the following table.

Score below	Number of students	Percentage of students
35	10	$\frac{10}{200} \times 100\% = 5\%$
50	50	$\frac{50}{200} \times 100\% = 25\%$
60	120	$\frac{120}{200} \times 100\% = 60\%$
75	195	$\frac{195}{200} \times 100\% = 97.5\%$

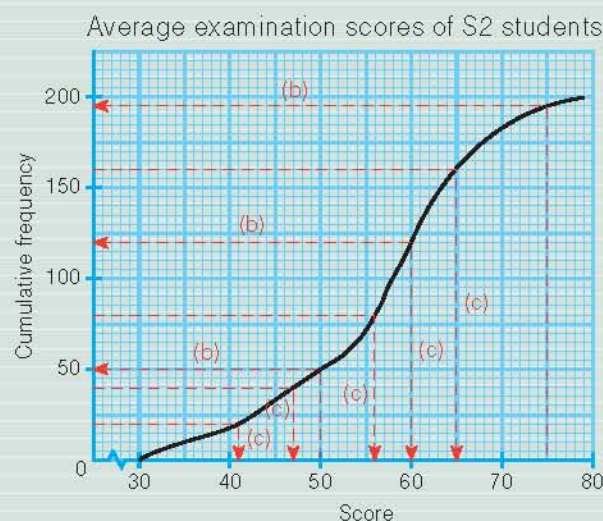


Figure I

(c) Referring to Figure I, complete the following table.

Percentage of students whose scores are below x	10%	20%	40%	60%	80%
Number of students whose scores are below x	20	40	80	120	160
Value of x	41	47	56	60	65

Now I see ...

From a cumulative frequency polygon or curve, we can find a value below which there is a given percentage of data.



From part (c) in Class Activity 7.2, when the percentage of students who scored below x is 40%, the value of x is 56, i.e. '40% of S2 students have scores less than 56'.

The score of 56 is called the 40th **percentile** and is denoted by P_{40} . Similarly, we can find that $P_{30} = 53$ from Figure I in Class Activity 7.2.

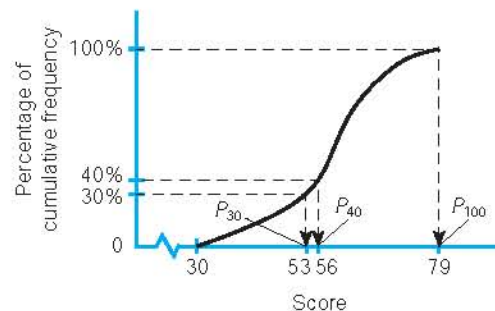


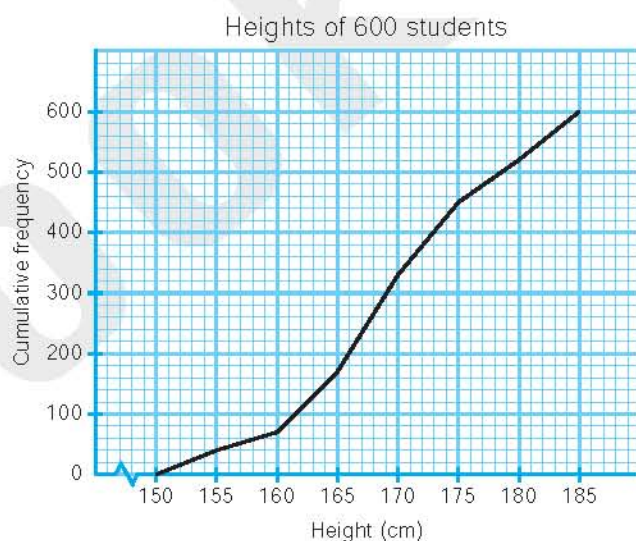
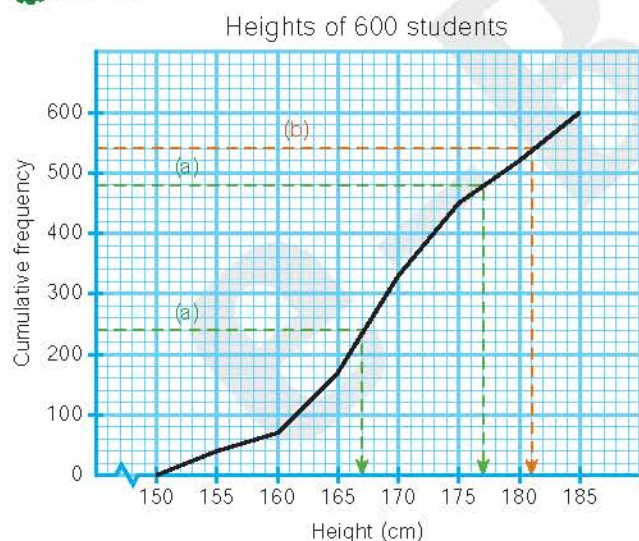
Figure 7.7

Example 7.6 Finding percentiles

The cumulative frequency polygon shows the heights of 600 students.

- Find P_{40} and P_{80} .
- If the tallest 10% of students are qualified to get into the school basketball team, find the height requirement for the school basketball team.

Solution



- Cumulative frequency corresponding to $P_{40} = 600 \times 40\% = 240$

From the figure, $P_{40} = \underline{167 \text{ cm}}$

Cumulative frequency corresponding to $P_{80} = 600 \times 80\% = 480$

From the figure, $P_{80} = \underline{177 \text{ cm}}$

From the figure, the total frequency is 600.

percentile 百分位數

- (b) [Analysis: The heights of the tallest 10% of students are between P_{90} and P_{100} .]

$$\begin{aligned}\text{Cumulative frequency corresponding to } P_{90} &= 600 \times 90\% \\ &= 540\end{aligned}$$

From the figure, $P_{90} = 181$ cm

\therefore The height requirement for the school basketball team is 181 cm or above.

Classwork 7.6

The cumulative frequency polygon shows the weights of 160 S2 students.

Weights of 160 S2 students



- (a) Find P_{20} and P_{70} .
 (b) If the heaviest 15% of students are required to have a body check, find the minimum weight requirement for a student to have a body check.

Quartiles

Class Activity 7.3

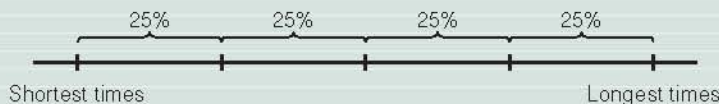
Aim: To learn the concept of quartiles

The following cumulative frequency polygon shows the time required for 200 children to complete a puzzle.

Time required for 200 children to complete a puzzle



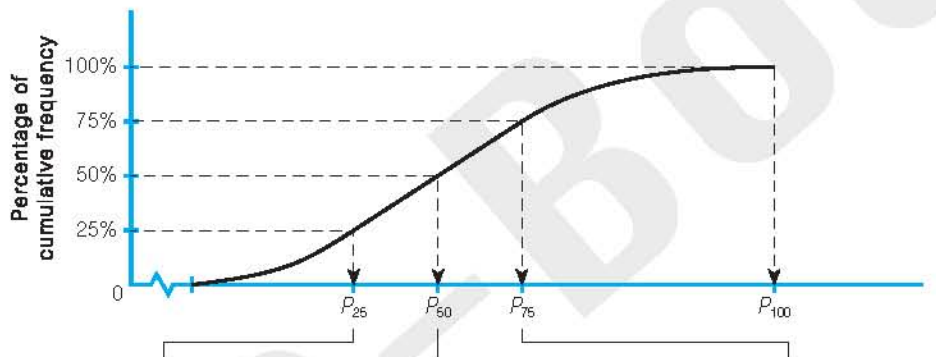
- Mark a time in the cumulative frequency polygon such that half of the children use less than the marked time to complete the puzzle.
- The first $\frac{1}{4}$ of children who completed the puzzle would receive a merit prize each. The children getting the prizes completed the puzzle within 105 min.
- As shown in the figure, if the time required for 200 children to complete the puzzle is rearranged in ascending order and then separated into 4 equal parts, each part contains 25% of children.



Which three percentiles can divide the number of data into 4 equal parts?

P_{25} , P_{50} and P_{75}

Consider a cumulative frequency curve (or a cumulative frequency polygon).



The 25th percentile, i.e. P_{25} , is also called the **lower quartile**, of which 25% of the data are less than this value.

The 50th percentile, i.e. P_{50} , is also called the **median**, of which 50% of the data are less than this value.

The 75th percentile, i.e. P_{75} , is also called the **upper quartile**, of which 75% of the data are less than this value.

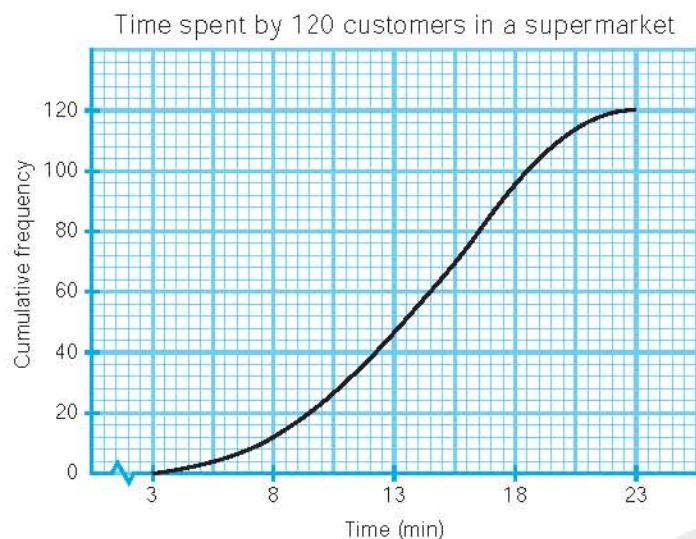
Figure 7.8

The percentiles P_{25} , P_{50} and P_{75} divide the number of data into 4 equal parts, so they are called **quartiles** and they are usually denoted by Q_1 , Q_2 and Q_3 respectively.

Note: We can also find Q_1 , Q_2 and Q_3 by the following method. When the total frequency is N , then the values corresponding to cumulative frequencies $\frac{N}{4}$, $\frac{N}{2}$ and $\frac{3N}{4}$ are Q_1 , Q_2 and Q_3 respectively.

Extension 7.2

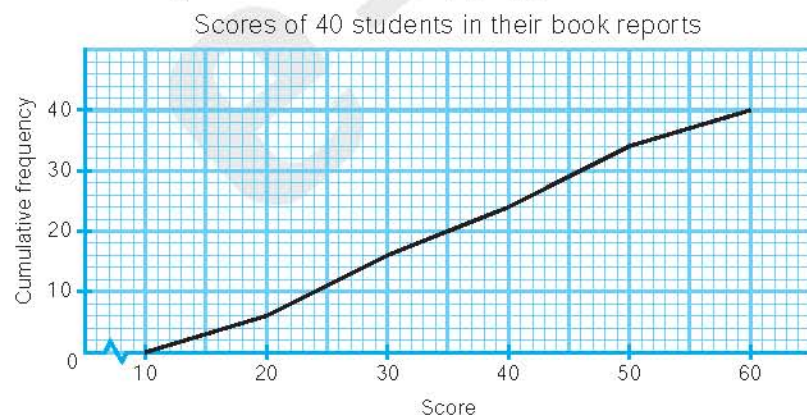
The following cumulative frequency polygon shows the time spent by 120 customers in a supermarket.



- Find the lower quartile.
- Find the median.
- Find the upper quartile.

Skills Upgrading Corner 7.3

The following cumulative frequency polygon shows the scores of 40 students in their book reports.



- Find the highest and lowest scores of 40 students in their book reports.
- Find
 - the lower quartile.
 - the median.
 - the upper quartile.
- If the bottom 20% of students have to do the book report again, at least what score should a student get to avoid doing a book report again?

Exercise 7C

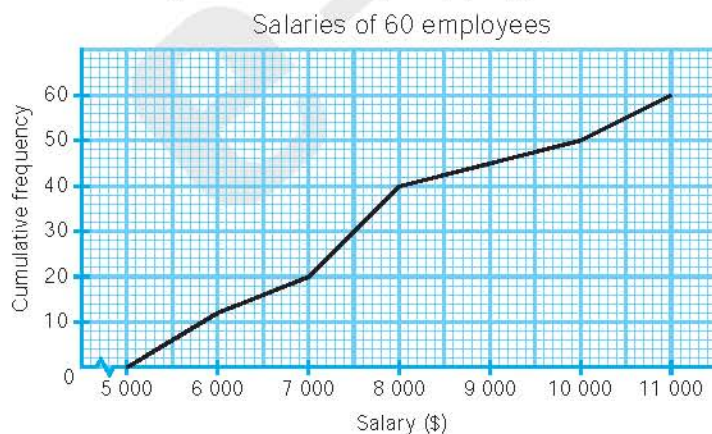
[Graph paper is provided in the Appendix.]

Level 1

- The following cumulative frequency polygon shows the ages of 100 residents in Tsuen Wan. Find P_{30} , P_{60} and P_{90} .



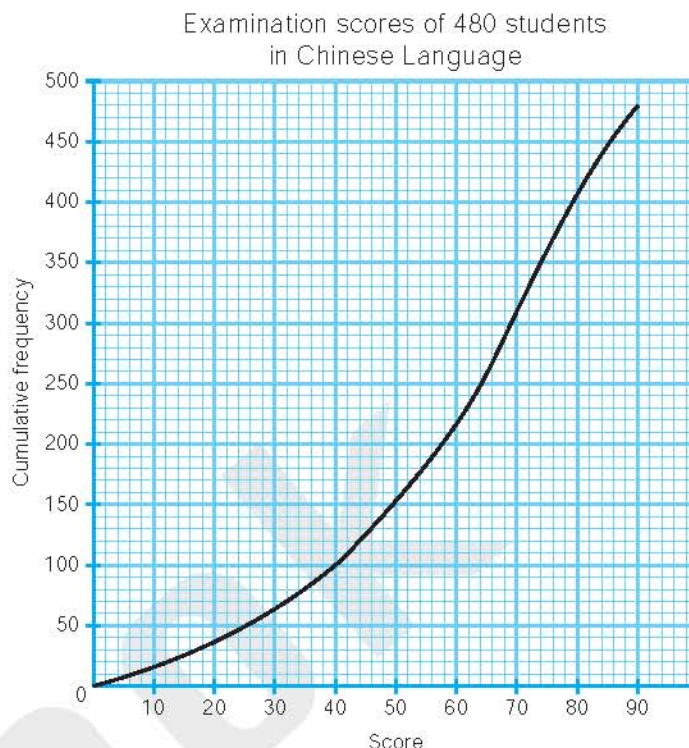
- The following cumulative frequency polygon shows the salaries of 60 employees.



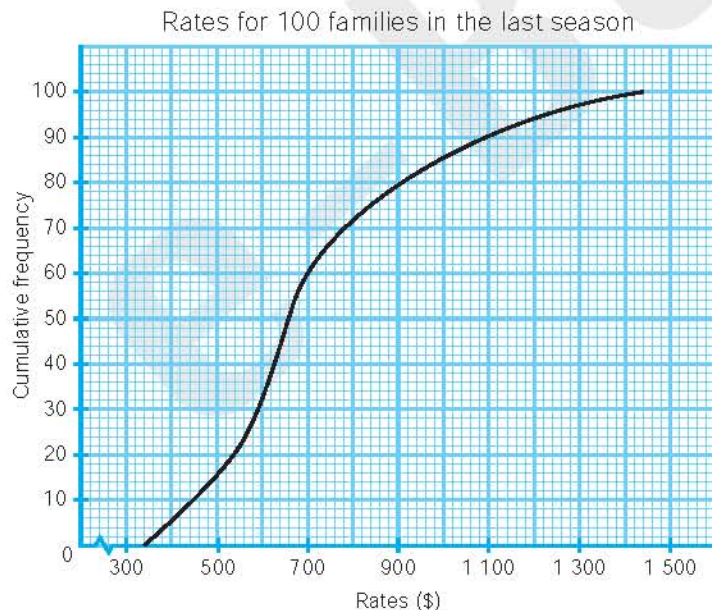
- Find P_{20} , P_{60} and P_{95} .
- Find
 - the lower quartile.
 - the median.
 - the upper quartile.

3. The cumulative frequency curve shows the examination scores of 480 students in Chinese Language.

- (a) Find
- the lower quartile.
 - the median.
 - the upper quartile.
- (b) If the top $\frac{1}{4}$ of students will receive a prize each, what is the minimum score of the students receiving prizes?



4. The following cumulative frequency curve shows the rates for 100 families in the last season.



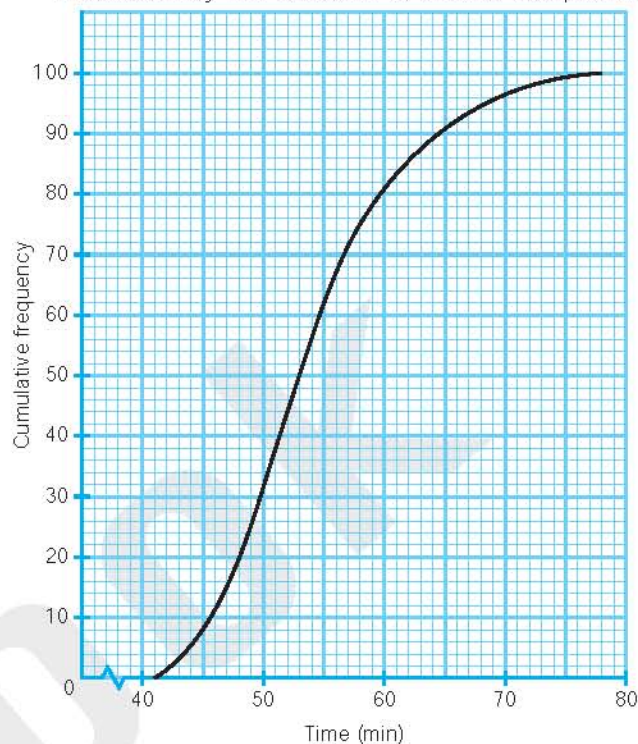
- (a) Find the maximum and minimum rates for 100 families in the last season.
- (b) Find
- the lower quartile.
 - the median.
 - the upper quartile.
- (c) How many families whose rates are between \$700 and \$1 100?
- (d) How many families whose rates are \$1 200 or above?

Level 2

5. The cumulative frequency curve shows the time taken by 100 students to finish a composition.

- (a) Find
 - (i) the lower quartile.
 - (ii) the median.
 - (iii) the upper quartile.
- (b) If finishing a composition within 45 min to 55 min is under a normal writing speed, how many students are writing at a normal speed?
- (c) The last $\frac{1}{4}$ of students finishing the composition will need to attend a writing workshop. Freddy has taken 57 min to finish a composition, does he need to attend the workshop? Explain briefly.

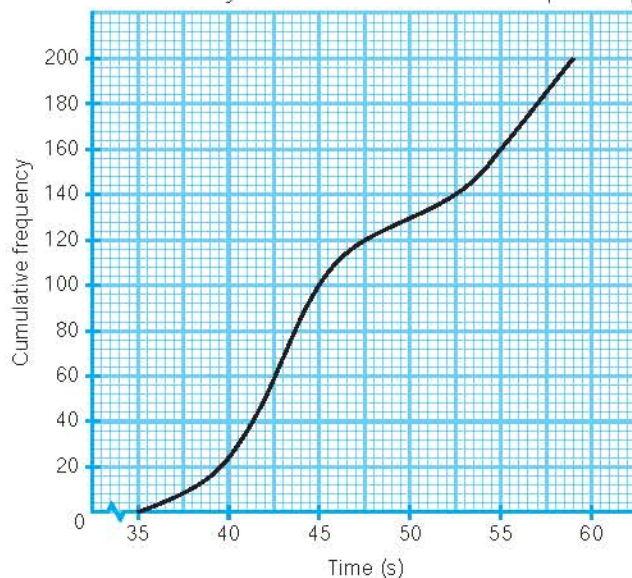
Time taken by 100 students to finish a composition



6. The cumulative frequency curve shows the time taken by 200 students to finish 20 push-ups.

- (a) Find
 - (i) the lower quartile.
 - (ii) the median.
 - (iii) the upper quartile.
- (b) If a student finishes 20 push-ups under 40 s will obtain a grade A, how many students will obtain grade A?
- (c) If a student taking 55 s or more to finish 20 push-ups will fail, how many students will fail?
- (d) If the time taken by Toby was P_{34} , how many students took more time than her to finish the push-ups? How long did Toby take?

Time taken by 200 students to finish 20 push-ups



7. The following table shows the test scores of 50 students in English Language.

Score	Frequency
20 - 29	3
30 - 39	8
40 - 49	12
50 - 59	17
60 - 69	6
70 - 79	3
80 - 89	1

(a) Construct a cumulative frequency table.

(b) Draw a cumulative frequency polygon.

[Horizontal scale: 10 squares = score of 10; vertical scale: 5 squares = 5 students.]

(c) From the graph in (b), find

(i) P_{10} and P_{90} .

(ii) the lower quartile, median and upper quartile.



Chapter Summary

Term Introduced

[This is a quiz to check your understanding of some special terms in this chapter. Match items in column A to column B appropriately.]

Column A

1. Frequency polygon •
2. Frequency curve •
3. Cumulative frequency polygon •
4. Cumulative frequency curve •
5. n th percentile •

Column B

- (a) The graph obtained by smoothing out a frequency polygon.
- (b) The graph using connected line segments to present the information of a cumulative frequency table.
- (c) There are $n\%$ of the data below a given value.
- (d) The graph using connected line segments to present frequency distribution.
- (e) The graph obtained by smoothing out a cumulative frequency polygon.



Check Yourself

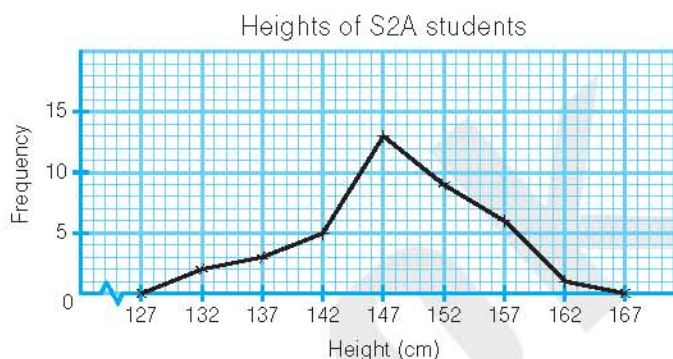
[This is a quiz to remind you of the basic concepts you have learned in this chapter. Each question tests a concept under the section listed on the right. Failure in any part of a question indicates a need to do a revision on the section listed.]

Section

1. The frequency polygon shows the heights of S2A students.

7.1

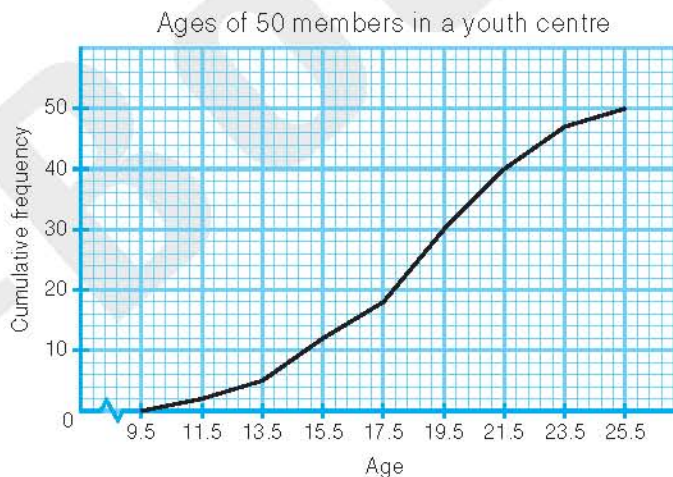
- (a) There are _____ students in S2A.
(b) Find the number of students whose heights are 159.5 cm or above.



2. The cumulative frequency polygon shows the ages of 50 members in a youth centre.

7.2

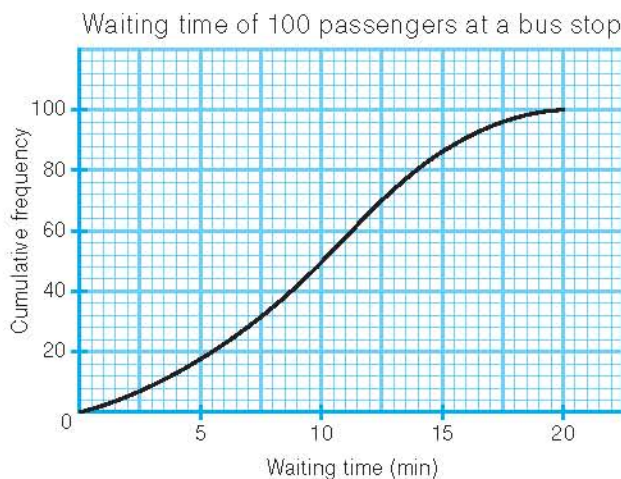
- (a) _____ members are younger than 15.5 years old.
(b) What percentage of members are aged 17.5 or above?



3. The cumulative frequency curve shows the waiting time of 100 passengers at a bus stop.

7.3

- (a) Median = _____ min
(b) For the passengers corresponding to the longest 30% of the waiting time at the bus stop, at least how long have they been waiting for?



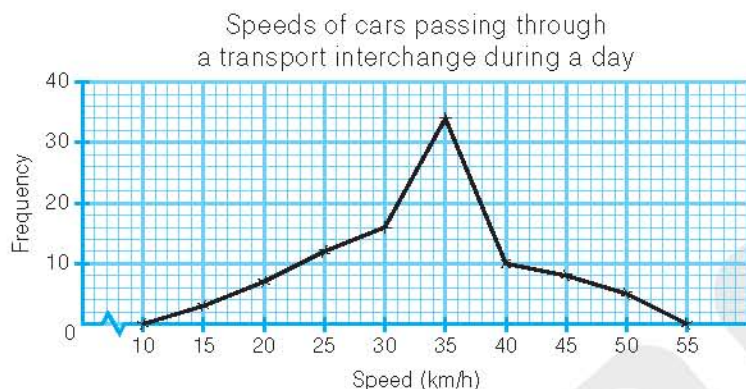


Revision Exercise 7

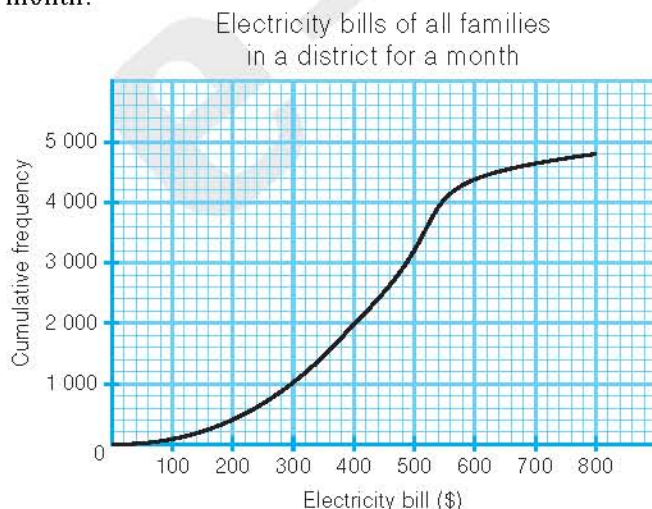
[Graph paper is provided in the Appendix.]

Level 1

1. The following frequency polygon shows the speeds of cars passing through a *transport interchange* during a day. It is given that the first class interval is 13 km/h - 17 km/h.



- How many cars passed through the transport interchange?
 - Which class interval has most of the cars?
 - How many cars travelled at a speed between 17.5 km/h and 22.5 km/h?
 - What percentage of the cars travelled at a speed between 27.5 km/h and 42.5 km/h? (Give your answer correct to 3 significant figures.)
2. The following cumulative frequency curve shows the electricity bills of all families in a district for a month.

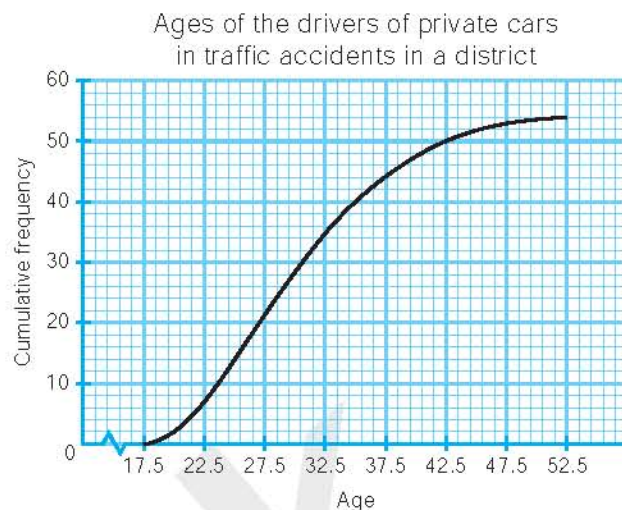


- Find the number of families in this district.
- Find the median electricity bill.
- Find P_{25} and P_{75} .
- Find the number of families with electricity bills between \$200 and \$600.

transport interchange 交通交匯處

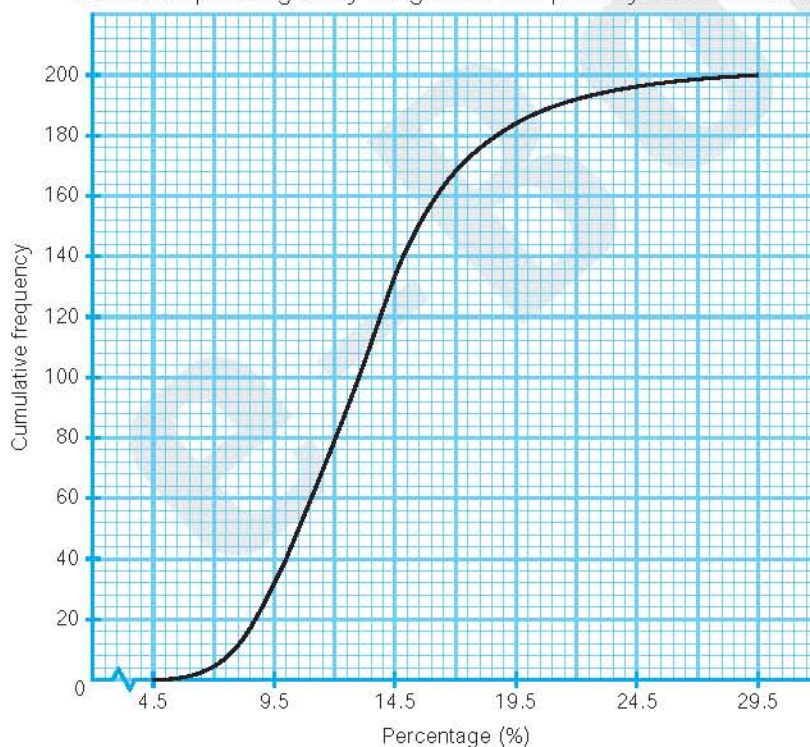
3. The cumulative frequency curve shows the ages of the drivers of private cars in traffic accidents in a district.

- How many drivers are shown in the diagram?
- Find the median age of the drivers.
- Find the percentage of the drivers whose ages are 39 or above. (Give your answer correct to 3 significant figures.)



4. The following cumulative frequency curve shows the weights of school bags as the percentages of the corresponding body weights of 200 primary school students.

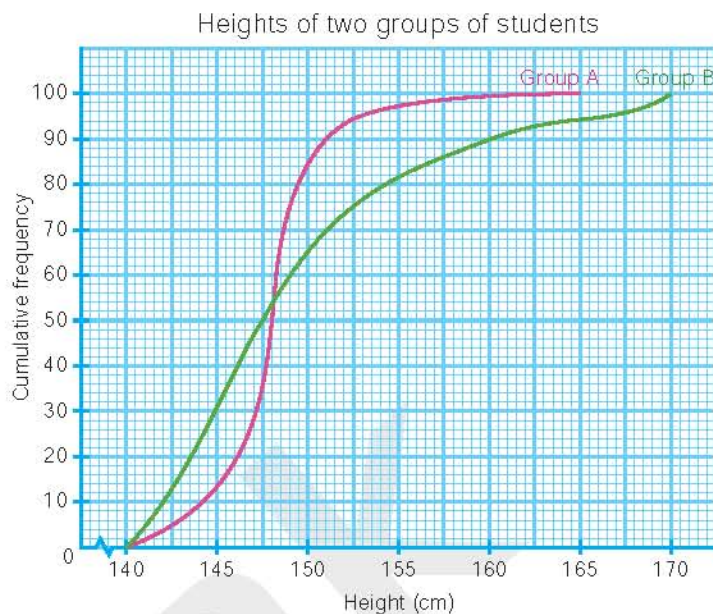
Weights of school bags as the percentages of the corresponding body weights of 200 primary school students



- Find
 - the lower quartile.
 - the median.
 - the upper quartile.
- According to a research, the weight of a school bag should not exceed 10% of the body weight. Among these 200 students, how many school bags are overweight?
- If Gigi weighs 23 kg and the weight of her school bag as the percentage of her body weight is the median, find the weight of her school bag.

5. The cumulative frequency curves show the heights of two groups of students.

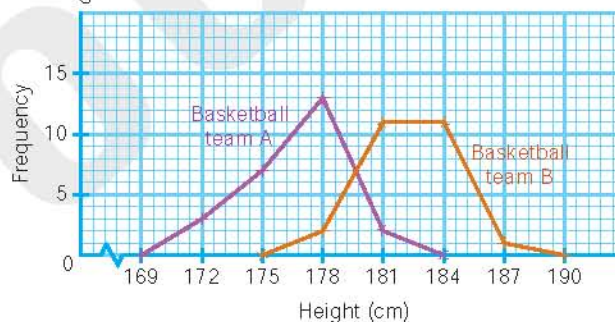
- Find the number of students in each group.
- For each group of students, find
 - the lower quartile.
 - the median.
 - the upper quartile.
- Which group has a larger difference between the upper quartile and lower quartile?
- For each group, if the tallest $\frac{1}{4}$ of students are chosen to join the swimming class, find the minimum height requirement for joining the swimming class.



6. The frequency polygons show the heights of the members in basketball teams A and B.

- Find the number of members in each basketball team.
- Which team has taller members in general?
- Which team do you expect to win the game? Why?

Heights of the members in basketball teams A and B



Level 2

7. The following are the weights of 40 people, correct to the nearest kg.

49 57 55 48 62 45 43 62 56 65 48 55 66 55 67 58 53 62 51 47
60 58 68 52 61 69 51 50 47 46 56 61 52 50 63 53 56 67 50 57

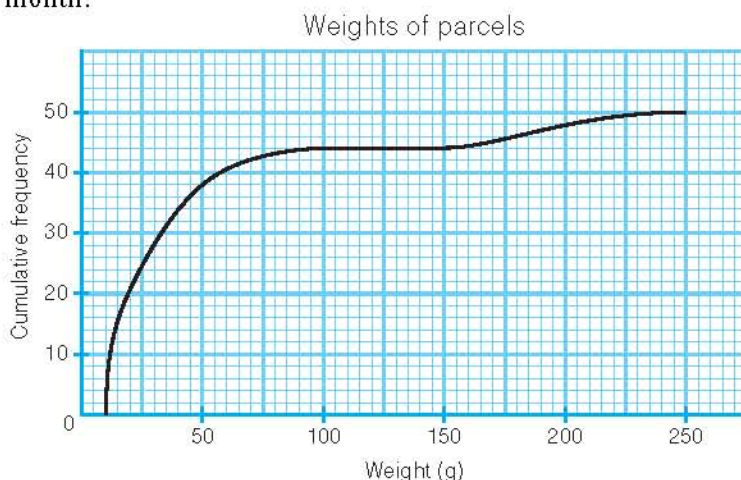
- Construct a frequency distribution table using 40 kg - 49 kg as the first class interval.
 - Draw a frequency polygon.
8. The following table shows the satisfactory level (with a full marks of 100) of 200 citizens towards their living in a city.

Satisfactory level	21 - 30	31 - 40	41 - 50	51 - 60	61 - 70	71 - 80	81 - 90	91 - 100
Frequency	1	3	8	22	73	47	25	21

- Construct a cumulative frequency table.
- Draw a cumulative frequency polygon.

[Horizontal scale: 5 squares = 10; vertical scale: 10 squares = 20 citizens.]

9. The following cumulative frequency curve shows the weights of the parcels sent by a company in a month.



- Find the maximum and minimum weights of the parcels sent by the company in that month.
 - Are there any parcels weighing 115 g?
 - If postage of a parcel weighing 100 g or below is the minimum charge, what percentage of the parcels whose postages are not the minimum charges?
10. The body mass index (BMI) is the most commonly used index to measure the body figure. The following cumulative frequency curve shows the BMI of 60 ladies.

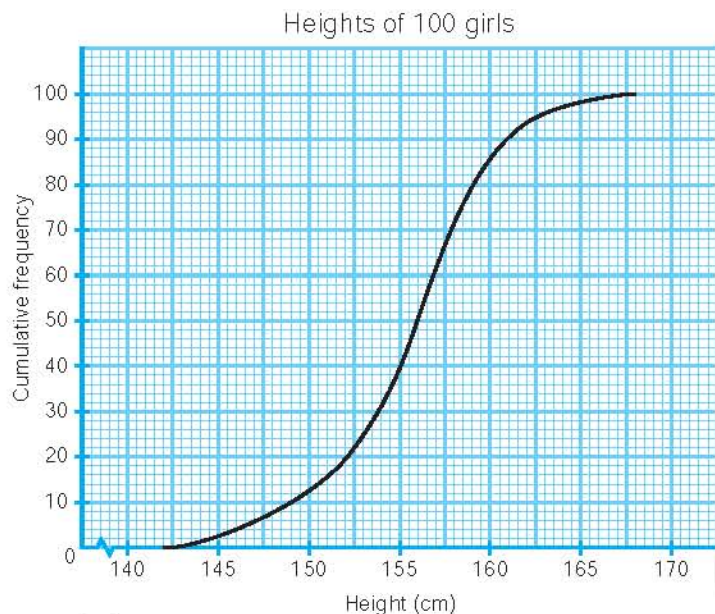


- A health centre has recorded the diet patterns of $\frac{1}{4}$ of the ladies with the smallest BMI. If the BMI of Amy is 16, does the health centre record her diet pattern? Why?
- The table shows the BMI classification.

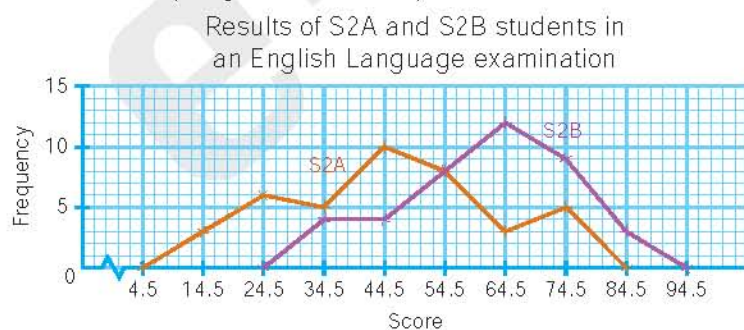
Classification	Underweight	Normal weight	Overweight	Extremely fat
BMI	Less than 18.5	Between 18.5 and 22.9	Between 23.0 and 24.9	Greater than 24.9

- Find the percentage of ladies who are extremely fat.
 - Does the median BMI of these 60 ladies fall into the classification of 'normal weight'?
- Describe the classification of the 60 ladies in general.

11. It is given that there are 100 girls applying to join the volleyball team in a school. The following cumulative frequency curve shows the heights of these 100 girls.



- (a) Find
- the lower quartile.
 - the median.
 - the upper quartile.
- (b) If the minimum height requirement for joining the volleyball team is 160 cm, find the number of girls who can join the team.
- (c) Due to insufficient members, the minimum height requirement is reduced to a level to allow additional 14 girls to join the team. What is the new minimum height requirement?
12. The following frequency polygons show the results of S2A and S2B students in an English Language examination (the pass mark is 50).



- (a) If the first class interval of the frequency polygon for S2A is 10 - 19,
- which class interval has the most students for S2A?
 - which class interval has the most students for S2B?
- (b) (i) Find the percentage of S2A students passing the examination.
(ii) Find the percentage of S2B students passing the examination.
- (c) (i) Construct a frequency distribution table to represent the combined results of S2A and S2B students in the English Language examination using 10 - 19 as the first class interval.
(ii) Which class interval has the most students?
(iii) Find the percentage of all students from the two classes passing the examination.

13. The following table shows the results of 50 children in a psychological test.

Score	21 - 35	36 - 50	51 - 65	66 - 80	81 - 95
Frequency	2	11	18	15	4

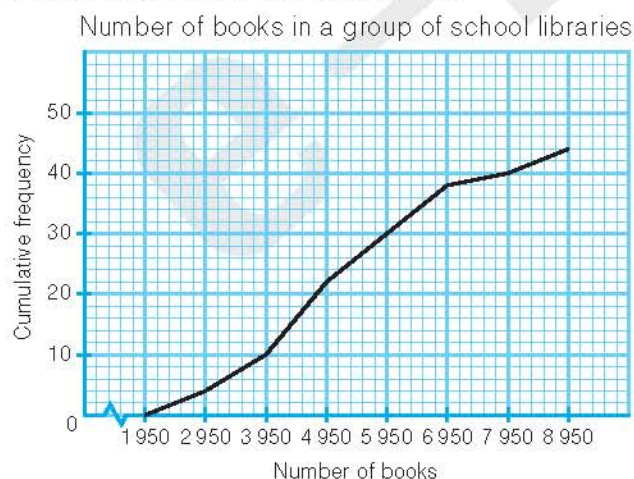
- Construct a cumulative frequency table.
- Draw a cumulative frequency polygon.
[Horizontal scale: 15 squares = score of 15; vertical scale: 10 squares = 10 children.]
- From the graph in (b), if any child scores 70.5 or above in the test prefers outdoor activities, what percentage of children prefer outdoor activities?

14. The following table shows the reward points of 80 credit card owners.

Reward point	1 - 1 000	1 001 - 2 000	2 001 - 3 000	3 001 - 4 000	4 001 - 5 000	5 001 - 6 000
Frequency	2	14	26	31	5	2

- Construct a cumulative frequency table.
 - Draw a cumulative frequency polygon.
[Horizontal scale: 10 squares = 1 000 reward points; vertical scale: 10 squares = 10 owners.]
- From the graph in (a)(ii),
 - if 500 reward points can exchange for a \$50 cash coupon of a supermarket, how many credit card owners can exchange for \$50 cash coupons?
 - if only 3 credit card owners have enough reward points to exchange for an electric boiler each, find the minimum reward points for exchanging an electric boiler.

15. The following cumulative frequency polygon shows the number of books in a group of school libraries.



- (a) From the above graph, complete the following table.

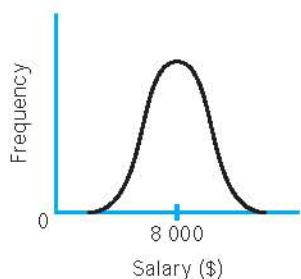
Number of books less than	1 950							
Cumulative frequency	0							

- Construct a frequency distribution table using 2 000 - 2 900 as the first class interval.
- Draw a frequency polygon.

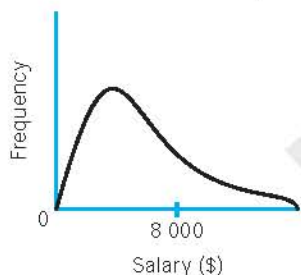
MC Question

16. It is known that Successful Company has 100 employees where most of them have salaries lower than \$8 000. Which of the following frequency curves can most probably represent the salaries of the employees?

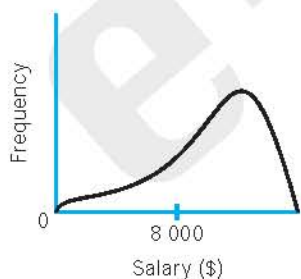
A. Salaries of employees in Successful Company



B. Salaries of employees in Successful Company



C. Salaries of employees in Successful Company

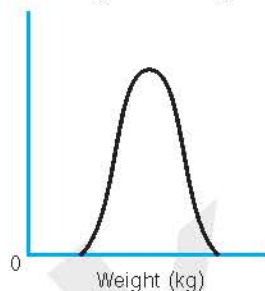


D. Salaries of employees in Successful Company

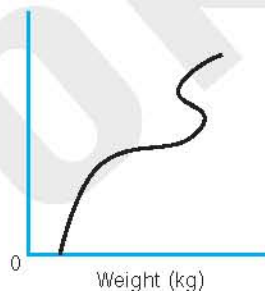


17. Which of the following can be a cumulative frequency curve?

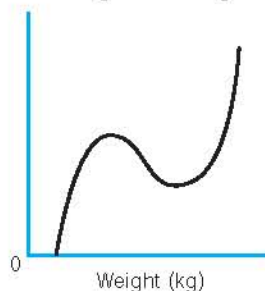
A. Weights of dogs



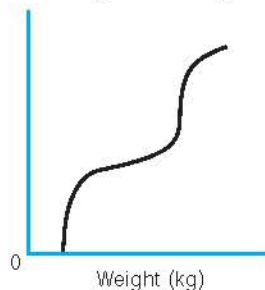
B. Weights of dogs



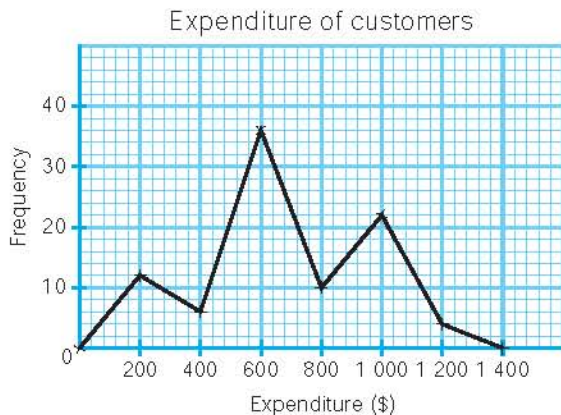
C. Weights of dogs



D. Weights of dogs



18. The following frequency polygon shows the expenditure of customers in a shopping centre last Sunday.

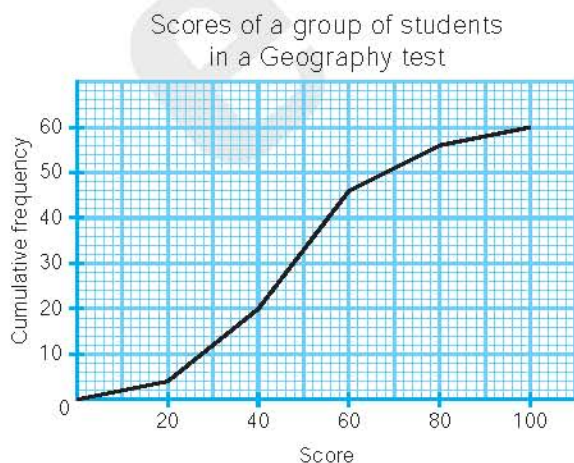


For any purchase of \$500 or above, a lucky draw ticket would be given. Find the percentage of customers who got lucky draw tickets last Sunday.

- A. $22\frac{2}{9}\%$
 B. 40%
 C. 72%
 D. 80%



19. The following frequency polygon shows the scores of a group of students in a Geography test.

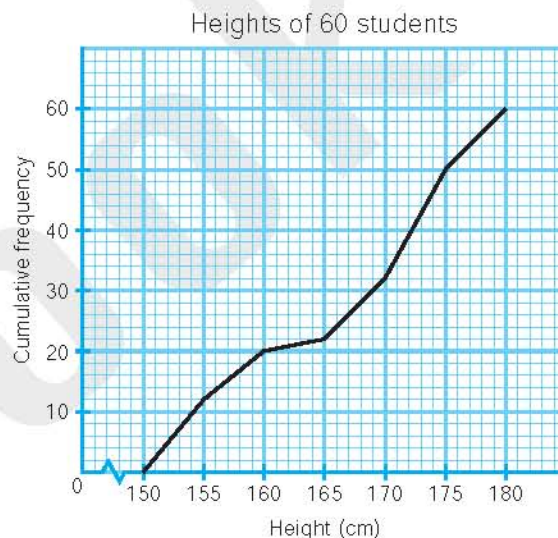


If students with scores of 80 or above will get grade A, find the number of students who get grade A.

- A. 2
 B. 3
 C. 4
 D. 5



According to the following figure, answer questions 20 – 21.



20. Find the median height of this group of students.

- A. 165 cm
 B. 168 cm
 C. 169 cm
 D. 175 cm



21. Find P_{60} .

- A. 171 cm
 B. 172 cm
 C. 175 cm
 D. 180 cm





Problem-solving and Exploring



Hint for the Title Page Question

- (a) According to the statistical graph on the title page, complete the following table.

Savings (\$)	Class boundaries (\$)	Frequency
1 - 5	0.5 - 5.5	
6 - 10	5.5 - 10.5	
11 - 15		
16 - 20		
21 - 25		
26 - 30		

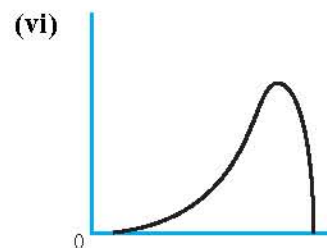
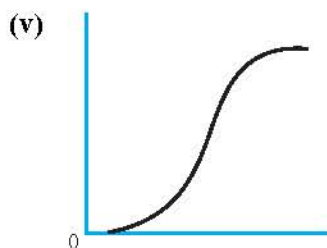
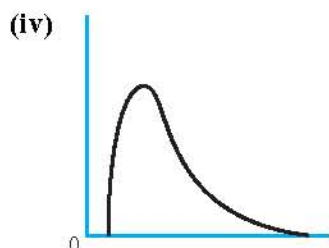
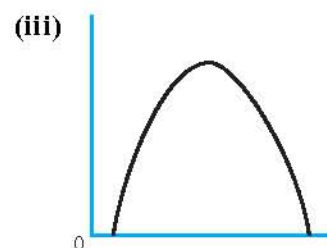
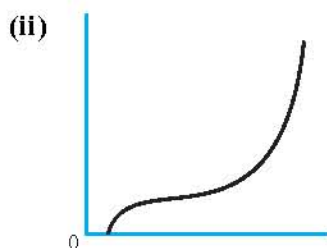
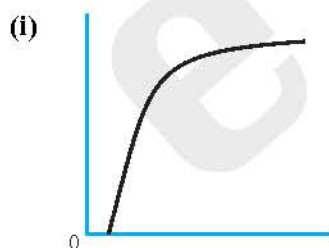


- (b) Estimate the maximum total savings of Minnie in that month.
 (c) Do you agree with Minnie? Explain briefly.



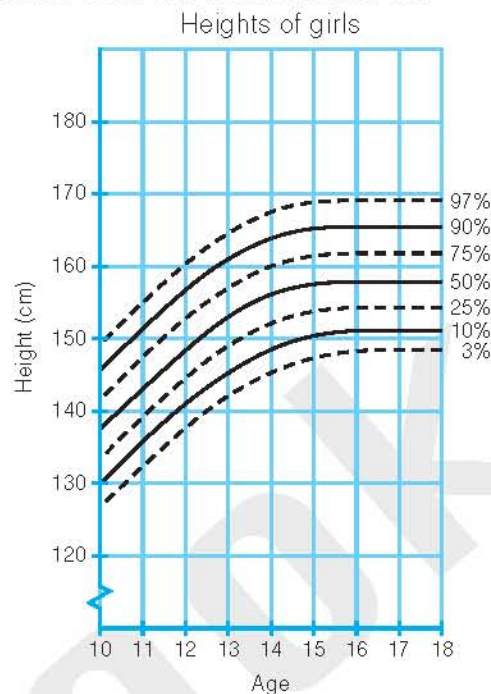
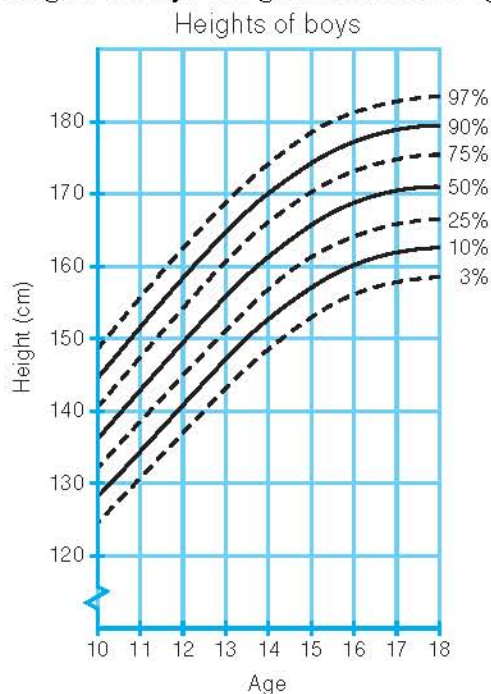
Additional Question

1. The following are three frequency curves and their corresponding cumulative frequency curves drawn by Jason. However, he has forgotten to write down the titles of the graphs.



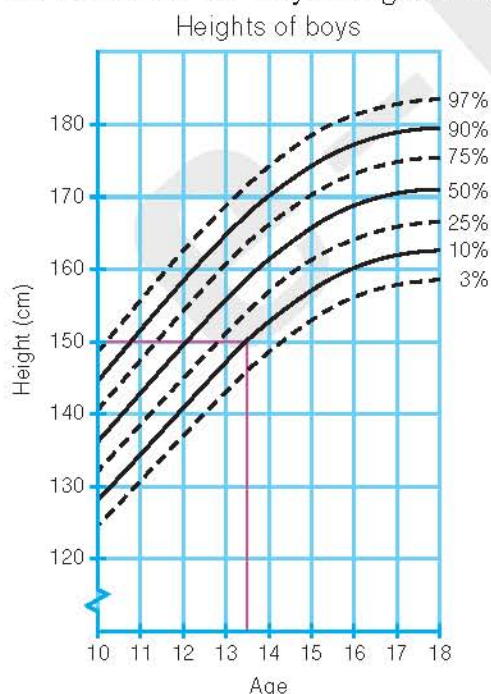
- (a) Determine which of the above graphs is a frequency curve or a cumulative frequency curve. Explain briefly.
 (b) Match each frequency curve with its corresponding cumulative frequency curve.

2. According to a survey on the growth of children in South China, the graphs of the percentiles of the heights of boys and girls at different ages from 10 to 18 are shown as follows.



[Source of Information: <http://www.cuhk.edu.hk/proj/growthstd/index.htm>]

John is 13.5 years old and his height is 150 cm. Based on Figure III, his height is the 10th percentile, i.e. John is the 10th boy among 100 boys of the same age arranged in ascending order of height.



According to the information in Figure I and Figure II, answer the following questions.

- Estimate the 90th percentile of the height of a boy/girl of 14 years old.
- Estimate the corresponding percentile of your height.



Construct a Cumulative Frequency Curve

Aim: To construct the cumulative frequency curve in Figure 7.6 (page 7.15) of Section 7.2

Required software: Microsoft Excel

Procedures:

Step 1: Input the information as shown in the figure.



	A	B	C	D	E	F	G
1	The weights of 40 students (correct to the nearest kg)					Weight less than (kg)	Cumulative Frequency
2	74	66	51	63		45.5	
3	46	52	57	60		50.5	
4	71	68	53	54		55.5	
5	65	70	57	59		60.5	
6	56	59	60	62		65.5	
7	54	49	69	67		70.5	
8	67	68	60	60		75.5	
9	54	64	70	62			
10	61	62	63	62			
11	62	75	65	61			

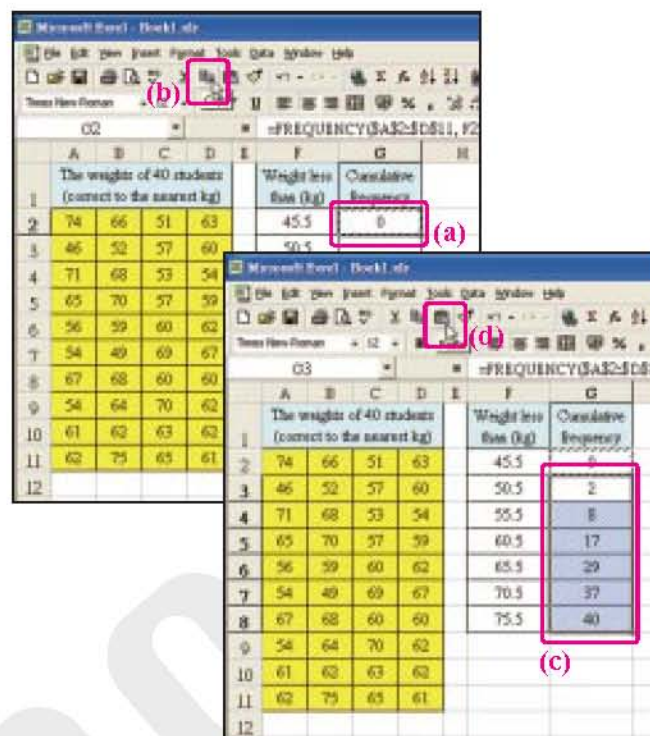
Step 2: (a) Select cell G2.


(b) Type in a formula '=FREQUENCY(\$A\$2:\$D\$11, F2)' and then press

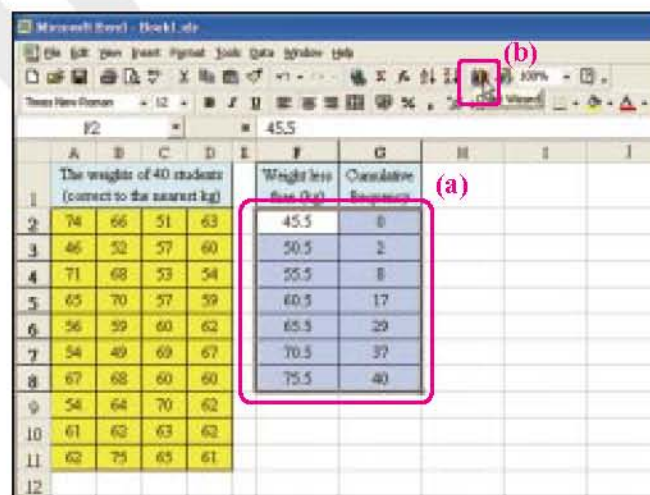
Enter.

	A	B	C	D	E	F	G
1	The weights of 40 students (correct to the nearest kg)					Weight less than (kg)	Cumulative Frequency
2	74	66	51	63		45.5	=FREQUENCY(\$A\$2:\$D\$11, F2)
3	46	52	57	60		50.5	
4	71	68	53	54		55.5	
5	65	70	57	59		60.5	
6	56	59	60	62		65.5	
7	54	49	69	67		70.5	
8	67	68	60	60		75.5	
9	54	64	70	62			
10	61	62	63	62			
11	62	75	65	61			

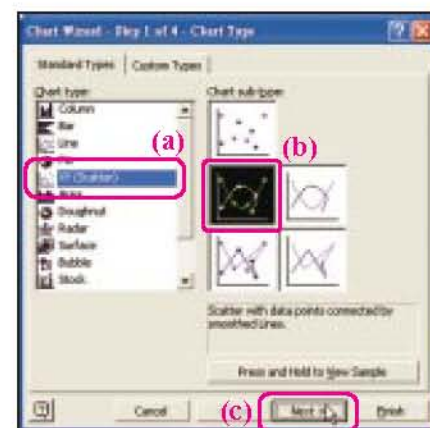
- Step 3:** (a) Select cell G2.
 (b) Click the icon .
 (c) Select the cells from G3 to G8.
 (d) Click the icon .



- Step 4:** (a) Select the cells from F2 to G8.
 (b) Click the icon .

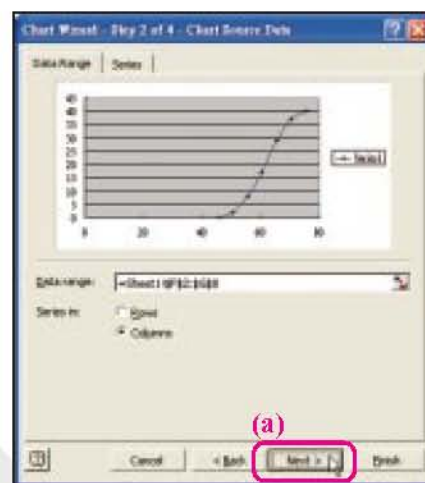


- Step 5:** For [Chart Type]
 (a) Select [XY (Scatter)] under [Chart type].
 (b) Select [Scatter with data points connected by smoothed lines] under [Chart sub-type].
 (c) Click the [Next] button.



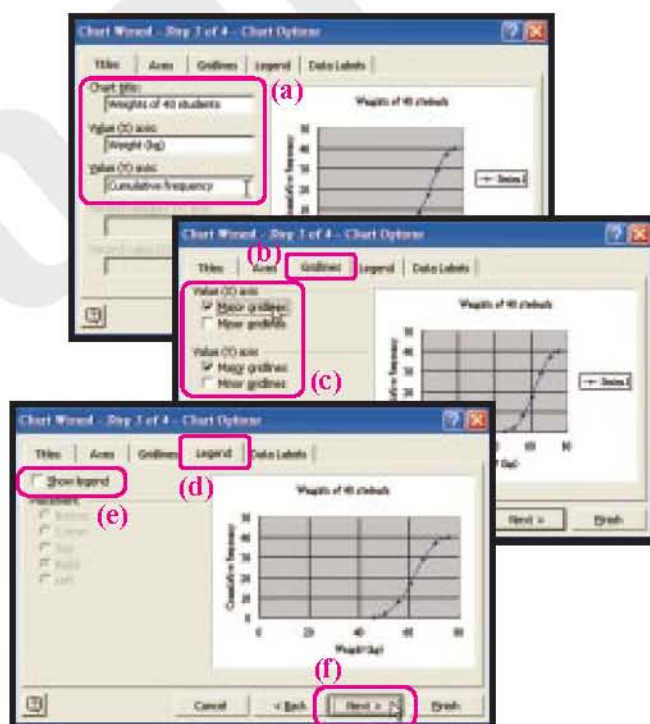
Step 6: For [Chart Source Data]

- (a) Click the [Next] button.



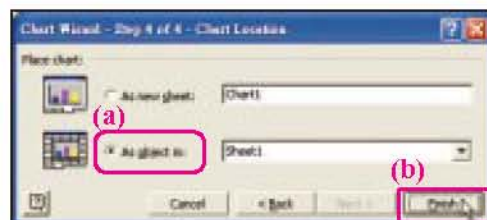
Step 7: For [Chart Options]

- (a) Fill in the box under [Chart title] with 'Weights of 40 students'; the box under [Value (X) axis] with 'Weight (kg)' and the box under [Value (Y) axis] with 'Cumulative frequency'.
- (b) Select [Gridlines].
- (c) Mark the boxes [Major gridlines] for both [Value (X) axis] and [Value (Y) axis].
- (d) Select [Legend].
- (e) Unmark the box [Show legend] to hide the legend.
- (f) Click the [Next] button.

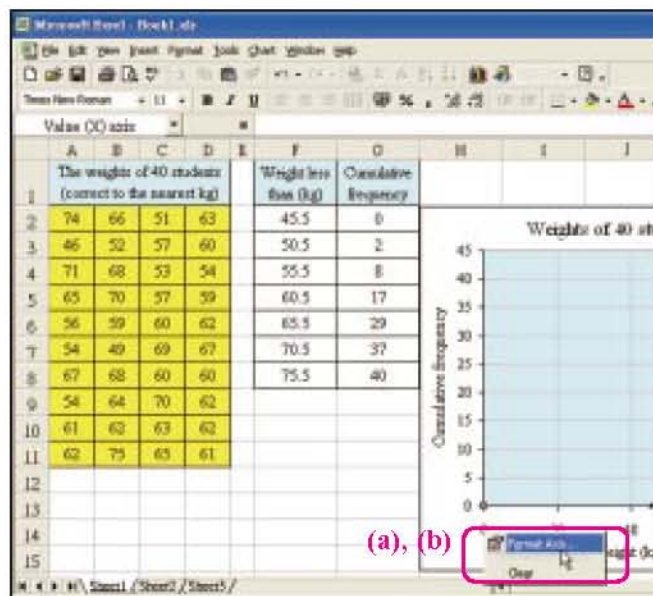


Step 8: For [Chart Location]

- (a) Select [As object in].
- (b) Click the [Finish] button.



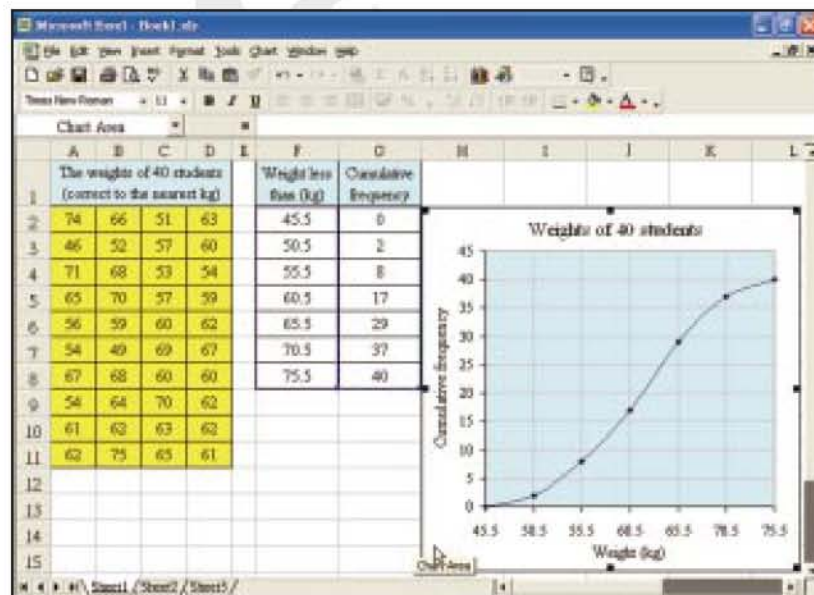
- Step 9:** (a) Move the cursor to any value on the horizontal axis and then click the right button on the mouse.
 (b) Select [Format Axis].



Step 10: For [Format Axis]

- (a) Select [Scale].
 (b) Fill in the box for [Minimum] with '45.5'; the box for [Maximum] with '75.5'; and the box for [Major unit] with '5'.
 (c) Click the [OK] button.

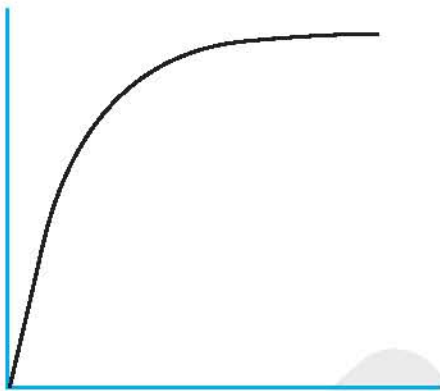
After completing the above steps, the following cumulative frequency curve will be obtained.



Practice

1. According to the information from the above cumulative frequency curve, answer the following questions.
 - (a) How many students weigh less than 55.5 kg?
 - (b) How many students weigh between 65.5 kg and 70.5 kg?
2. Suppose that the data in cells B6, B7 and B8 change to 95, 94 and 86 respectively, how does the cumulative frequency curve change? Explain your answer briefly.
3. Change the data in the cells from A2 to D11 so that the cumulative frequency curve becomes the curves shown below. What is the characteristic of these data?

(a)



(b)

